PB83-106237

Air Quality Modeling and the Clean Air Act Recommendations to EPA on Dispersion Modeling for Regulatory Applications

American Meteorological Society
Boston, MA

Prepared for

Environmental Protection Agency
Washington, DC

1981

U.S. Department of Commerce
National Technical Information Service
NTIS

PB83-106237

AIR QUALITY MODELING and the CLEAN AIR ACT:

Recommendations to EPA on Dispersion Modeling for Regulatory Applications

AMERICAN METEOROLOGICAL SOCIETY

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA 22161

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO.

2\.

3. RECIPIENT'S ACCESSION NO.
PB83 106237

4. TITLE AND SUBTITLE
AIR QUALITY MODELING AND THE CLEAN AIR ACT:
Recommendations to EPA on Dispersion Modeling for Regulatory Applications

5. REPORT DATE
1981

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
American Meteorological Society
Boston, Massachusetts

10. PROGRAM ELEMENT NO.

11. CONTRACT/GRANT NO.

12. SPONSORING AGENCY NAME AND ADDRESS
U.S. Environmental Protection Agency

13. TYPE OF REPORT AND PERIOD COVERED

14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

Through the Clean Air Act and its amendments, the U.S. government endeavors to protect and enhance the quality of the nation's air resources by a variety of air quality management systems. Source emissions, meteorological observations, air quality models, and air qulaity data are the basic components of an air quality management system Air quality models provide a scientific means of relating emissions and atmospheric processes to provide estimates of ambient air quality values. The models consist of mathematical equations relating the release of effluents into the atmosphere to expected concentrations in ambient air. Air quality models, therefore, can be used to identify and to evaluate the level of controls required to solve industrial and urban air-pollution problems. This report addresses the use of air quality models in a regulatory framework and represents the effort of the American Meteorological Society (AMS) under the original Cooperative Agreement with the Environmental Protection Agency, September 6, 1979, through March 31, 1981.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Air Pollution Air Quality Models		EPA68A

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
	20. SECURITY CLASS *(This page)*	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

AIR QUALITY MODELING and the CLEAN AIR ACT:

Recommendations to EPA on Dispersion Modeling for Regulatory Applications

REPORT PREPARED UNDER A
COOPERATIVE AGREEMENT WITH THE
ENVIRONMENTAL PROTECTION AGENCY

AMERICAN METEOROLOGICAL SOCIETY
Boston, Massachusetts 1981

Available from

AMERICAN METEOROLOGICAL SOCIETY
45 Beacon Street
Boston, Massachusetts 02108

Preface

Like most other scientific specialities, atmospheric transport and diffusion modeling has gone through definable phases. The first, from about 1920 to 1947, was a period in which the basic theoretical concepts now used in air quality modeling were developed but during which experimental research was considered a necessary handmaiden to theory. L. F. Richardson and G. I. Taylor were key figures in the development of diffusion modeling, and one need only mention the close relationship between the first use of statistical expressions to represent turbulent diffusion by O. G. Sutton and the field experiments at Porton, England, to be reminded of the attitude of the scientists working in that period. Indeed, many of the transport and diffusion models used today are based on the Gaussian core suggested by Sutton 35 years ago. Much of this early work was directed toward the understanding of ground-level releases and diffusion processes near the source.

The period from 1948 through 1968 might be characterized as the experimental phase. There was an intense desire on the part of those working in turbulence and diffusion to improve their understanding of the physics of atmospheric processes. Elaborate facilities and experiments were established to further this understanding. A few important examples are Greenglow, Prairie Grass, the diffusion studies at Dugway, as well as long-term studies of plume behavior at the Tennessee Valley Authority and in the United Kingdom. During this period most of the U.S. efforts were supported by the Atomic Energy Commission, various branches of the Armed Services, the Public Health Service, and the Tennessee Valley Authority. There were also many studies of the properties of the atmospheric boundary layer, mainly conducted using data collected on tower installations at national laboratories and universities. Improvements in the theoretical treatment of turbulence and diffusion consisted largely of refinements of existing concepts, based on good experimental data. Most of these experiments and theories appear in the review document Meteorology and Atomic Energy (Slade, 1968) and in Atmospheric Science and Power Production (Randerson, 1981).

During the 1970's, fundamental knowledge on transport and diffusion continued to increase. Some examples of progress are the use of convective scaling to estimate diffusion in the full depth of the daytime convective boundary layer, the development of urban photochemical models, and the development of numerical simulation techniques for meteorological, chemical, and physical processes. In addition, physical modeling became widely used to study transport and diffusion over complex terrain. By contrast, regulatory needs defined by the Clean Air Act and its amendments have tended to encourage the application of specific models with fixed technical ability. All the applied models in the Environmental Protection Agency (EPA) UNAMAP model series have Gaussian cores and are primarily based on fundamental knowledge gained prior to 1970.

The 1970 amendments to the Clean Air Act produced significant motivation for the use of models. Diffusion models were needed to support the law. In fact, models were specifically mentioned in the 1977 amendments. As regulatory procedures developed after both amendments, models became a major component of enormously expensive and important decisions. Precise limits on emissions, and even decisions about acceptability of a new source in a given area, are often based on numbers provided by an official set of diffusion models. In an effort to ensure equity in this process, the Congress asked the EPA to provide a forum for model discussions and to develop regulations regarding their use.

During public hearings in 1977, leading to the EPA Guideline on Air Quality Models, it was suggested that the American Meteorological Society (AMS) could provide a service to EPA and to the modeling community by providing scientific advice on the evaluation of air quality models. Indeed, the AMS had issued position statements on stability classification, sigma curves, and model accuracy and had conducted a major workshop in the field. In late 1978, EPA invited formal discussions, and a Cooperative Agreement was consummated in September 1979. The AMS offered to provide expertise and assistance in evaluating technical aspects of air quality models that may be used for applications as required by the Clean Air Act amendments of 1977. In particular, the AMS agreed to:

1. Conduct a review of and make recommendations on previous work concerning air quality models accomplished by EPA,
2. Undertake a general review of the state of knowledge on air quality modeling,

3. Offer suggestions concerning recommended air quality models and criteria for their selection,

4. Evaluate data-base requirements for use with models, and

5. Assist EPA in the conduct of a National Conference on Air Quality Modeling to meet the requirements of the Clean Air Act. (This task has since been deleted.)

The AMS motivation for this effort was to provide a mechanism by which the meteorological community could more directly suggest, propose, and counsel on the use of sound meteorological principles of relevance to EPA's programs and policies that deal with air quality models and assessment and EPA's administration of them. We have refrained from becoming involved in specific rulemaking decisions associated with modeling. Rather, we have concentrated on establishing procedures for using new and scientifically sound techniques for modeling boundary-layer processes for application in a regulatory framework.

Early in the Cooperative Agreement it became clear to the Steering Committee that both EPA and AMS interests would be best served by meeting regularly (approximately every two months) to discuss and advise the EPA on their ongoing activities in model development and application. EPA has been working continuously to improve air quality modeling by pursuing suggestions made through various public comment opportunities. To engage too closely with them in this area would be potentially to violate constraints placed on us by the AMS Executive Committee regarding regulatory involvement. However, not to provide suggestions and assistance on technical matters would compromise the intent of our Cooperative Agreement. Thus we have, for example, not been involved in the drafting or the review of proposed new guidelines nor have we commented on any specific EPA modeling decisions. We have, however, felt it important to comment on the use of scientific information in a general sense, for example, the EPA posture on acceptance of new ideas and on some of the difficult problems confronting regulatory modeling. In this regard our report may not represent a totally objective, unbiased, purely scientific reflection of the state of the art in air quality modeling. Rather it represents a collection of thoughts on this subject, colored to a large extent by the ongoing debates about the usefulness and appropriate use of models and modeling in the management of our air resources.

This Cooperative Agreement, which has been under way for some 2 years, has been under the direction of a small Steering Committee consisting of two former chairmen of the AMS Committee on Turbulence and Diffusion (Darryl Randerson and Steven R. Hanna) and two former chairmen of the AMS Committee on Air Pollution (Douglas G. Fox and Bruce A. Egan). Fred D. White has served as the Staff Director on a part-time basis. L. E. Niemeyer, Herschel H. Slater, and Joseph A. Tikvart have been our principal EPA contacts.

The Steering Committee met 20 times during the course of this project. We have formally utilized the advice of over 40 individuals and hundreds of others on an informal basis. Their assistance and thoughtful suggestions are greatly appreciated.

Bruce A. Egan
Douglas G. Fox
Steven R. Hanna
Darryl Randerson
Fred D. White

Contents

1. Introduction and Recommendations 1

2. Review of Past Activity 6
   - 2.1 Introduction 6
   - 2.2 Specialists' Conference 8
   - 2.3 Hearings and National Modeling Conference (1977) 14
   - 2.4 Public Hearings on the 1980 Draft Guideline 15

3. State of the Science of Air Quality Modeling 18
   - 3.1 Scientific Basis for Modeling and Known Performance Characteristics 18
   - 3.2 Summaries of Review Papers on the State of the Science 19
   - 3.3 Research Needed to Improve Air Quality Models 22

4. The Evaluation of Air Quality Models for Regulatory Applications 33
   - 4.1 Introduction 33
   - 4.2 Solicitation of New Models 35
   - 4.3 Performance Evaluation 37
   - 4.4 Principles of Selection 41
   - 4.5 Current Approaches to Model Comparisons 43

5. Meteorological and Air Quality Data 48
   - 5.1 Meteorological Data 48
   - 5.2 Air Quality Data 49
   - 5.3 Data Representativeness 51

References 55

Appendixes

A. Summary Report on EPA's 1977 Air Quality Modeling Conferences, Gordon A. Beals 63
B. The Air Quality Modeling Conference of December 14 and 15, 1977, Gordon A. Beals 76
C. Comments on the Adequacy of EPA Responses to the Issues Raised in the Public Meetings and Air Quality Modeling Conference, Gordon A. Beals 88
D. Summary Report on EPA's 1980 Air Quality Modeling Conferences, Gordon A. Beals 100
E. Modeling Long Range Transport and Diffusion, Arthur Bass 136
F. Validation and Performance Criteria for Air Quality Models, Norman E. Bowne 159
G. Dispersion in the Vicinity of Buildings, R.P. Hosker, Jr. 172
H. Modeling Visibility, Douglas A. Latimer, Robert W. Bergstrom, Clark D. Johnson, and James P. Killus 188
I. Wet and Dry Removal, Resuspension, and Atmospheric Chemistry of SO_x and NO_x, D. Randerson 204
J. Transport and Diffusion Modeling and Its Application--1980, Maynard E. Smith 228
K. Judging Air Quality Model Performance, Douglas G. Fox 271
L. On-Site Meteorological Instrumentation Requirements to Characterize Diffusion from Point Sources--A Workshop, 15-17 January 1980, Raleigh, N.C., G. F. Hoffnagle M.E. Smith, T.V. Crawford, and T.J. Lockhart 282

1

Introduction and Recommendations

Through the Clean Air Act and its amendments, the United States Government endeavors to protect and enhance the quality of the nation's air resources by a variety of air quality management systems. Source emissions, meteorological observations, air quality models, and air quality data are the basic components of an air quality management system. Air quality models provide a scientific means of relating emissions and atmospheric processes to provide estimates of ambient air quality values. The models consist of mathematical equations relating the release of effluents into the atmosphere to expected concentrations in ambient air. Air quality models, therefore, can be used to identify and to evaluate the level of controls required to solve industrial and urban air-pollution problems. This report addresses the use of air quality models in a regulatory framework and represents the effort of the American Meteorological Society (AMS) under the original Cooperative Agreement with the Environmental Protection Agency (EPA), September 6, 1979, through March 31, 1981.

One of the key elements of air quality legislation and regulation in the United States is the prominent role of air quality modeling in decisions affecting the location of a facility and levels of emission controls. Both the National Ambient Air Quality Standards (NAAQS) and the Prevention of Significant Deterioration (PSD) increments are expressed in terms of ground-level pollutant concentrations. Atmospheric dispersion models provide the link between the yet-to-be-constructed or yet-to-be-modified source and potential pollutant concentrations at any receptor location. The PSD provisions place particular emphasis on such predictive modeling.

Models are a mathematical representation of reality. We have an incomplete understanding of the complex physical and chemical processes involved in the dispersion, transformation, and deposition of pollutants. Because of the complexity and inherent turbulent stochastic properties of the atmosphere, some limitations to the predictive capability of even the best model will always remain. Uncertainties in emissions data and meteorological data also contribute to an overall uncertainty of model predictions. Nevertheless, to the extent that models reflect our best understanding of the relevant physical processes, they represent a scientifically sound and objective means for taking into account differences of source geometries and of topographic and meteorological settings in assessing ambient air quality impacts. Thus, models provide a logical and environmentally equitable basis for decision-making.

The basic research that underlies air quality modeling includes studies of the structure of atmospheric boundary-layer turbulence, the specific process occurring within the boundary layer, and interaction of the boundary layer with the free atmosphere. At present, air quality models are distinct from weather-prediction and climate-prediction models, though each draws on basic research in boundary-layer physics. Progress in air quality modeling is likely to depend on progress in basic research and would benefit from a closer linkage with weather prediction and climate modeling. In the short run, the separation exists, and this report addresses specific steps to be taken now to improve air quality modeling.

Our conclusions and recommendations follow. Supporting discussion appears in the following chapters.

1. We recommend that EPA strongly endorse the use of atmospheric dispersion models for assistance in making decisions related to the management of air quality. Based on our review we have concluded that these models are useful and appropriate technical tools, but the interpretation of model predictions should include a consideration of statistical confidence intervals.

2. We recommend that EPA continue to enhance research and development activities to increase our understanding of the following topics (not listed in order of priority):

(a) Long-range transport and transformation of pollutants thought to be involved in acid precipitation;

(b) Dispersion in regions of mountainous terrain;

(c) Modeling of chemically reactive pollutants of concern to the formation of ozone and other important atmospheric materials;

(d) Modeling of the atmospheric transport, deposition, and ultimate fate of toxic materials in the ecosystem; and

(e) Identification of representative meteorological information from considerations of long-term off-site climatology and local short-term meteorology.

3. We recommend that EPA anticipate better the direction of regulatory modeling needs. Researchers must be more sensitive to the regulatory context of their products and are encouraged to provide a finished product in a timely fashion. Regulators must be willing to incorporate new improved procedures when they become available. To bring researchers and regulators together, we recommend jointly conducted workshops or other procedures, drawing heavily on scientists outside EPA, to improve the use of new scientific findings and to provide mutually acceptable direction to EPA programs.

4. We recommend that EPA adopt and issue specific guidance, for example, as a part of the PSD regulations, relating to the meteorological data needs described in Chapter 5 and Appendix L. The importance of using appropriate meteorological data input to air quality models is stressed. The suggestions in Chapter 5 and Appendix L focus on the importance of on-site measurements of turbulence intensities, wind velocity, temperature gradient, and mixing depth.

5. We recommend that EPA adopt a specific set of measures to evaluate the performance of new and existing models statistically. This evaluation will be based on the comparison of model simulations with observations. The statistical evaluation should be based on an identified set of performance measures. One suggestion for a set of performance measures is embodied in Chapter 4 and is based on results of a workshop conducted in September 1980 under the AMS/EPA Cooperative Agreement.

6. We recommend that EPA review all models, including those currently recommended in the modeling Guideline and those being considered for regulatory applications. This review should be based on a statistical evaluation using the performance measures recommended in 5 above and on a scientific review by an impartial panel of experts. This review is intended to evaluate the scientific merits of the model. We believe that the scientific review should have the following characteristics:

(a) Adequate financial and staff support should be provided to allow the conduct of the statistical evaluations and to ensure feedback of review findings to model developers;

(b) No less than five reviewers knowledgeable in the science of modeling and in the regulatory application of models should participate;

(c) Reviewers must be provided with the results of the statistical performance evaluation; and

(d) reviewers should consider all models in a given category (point source, urban, etc.).

We further recommend that a group distinct from those responsible for regulatory modeling within EPA assume overall responsibility for this review. AMS offers to provide an overview of this program.

7. We recommend that EPA refocus its efforts to achieve consistency in modeling applications. The focus should be shifted away from issues of specific numerical agreement between alternative models (and the use of a limited set of models) and toward issues of consistency of approach to modeling problems. We are concerned that EPA has overdrawn its mandate to specify models "with reasonable particularity" to imply numerical agreement. Instead, EPA should develop a review process that is consistent, for example, in terms of meteorological data requirements and types of dispersion model outputs used. This review process should also treat new models responsibly--recognizing the admonition to EPA contained in Judge Robinson's Alabama Power decision, that in questions of modeling controversy "EPA should move to adopt the more accurate procedure."

8. We recommend that EPA recognize differences between two numerical results as being "significant" only if they are outside a confidence interval statistically constructed about the predictions. For example, model components, such as plume rise or dispersion coefficients may be based on different theories or data sets but still may produce results that agree within the expected accuracy of the method.

9. We recommend that EPA consider adopting a procedure to evaluate compliance with the short-term standards on the basis of more robust statistics than the highest second-highest value. For example, consideration of a National Ambient Air Quality Standard (NAAQS) violation or Prevention of Significant Deterioration (PSD) increment consumption could be reformulated to the 95th percentile. This approach would necessitate an adjustment of the numerical value and/or the manner in which a standard is calculated to ensure that emission limitations in general would not be more or less stringent than they are at present. Such adjustments may differ for exceptional circumstances involving complex terrain settings. The basis for this recommendation is that more frequently occurring meteorological events can be predicted with more confidence than rarely occurring events. We suggest that not only would this change permit models to predict the more frequently occurring events more accurately, but it also could allow a reduction in the quantity of meteorological data currently required as input to models. The procedure would have the additional benefit of reducing the subjective judgmental aspects about what constitutes realistic worst-case conditions.

10. We recommend that EPA call for the use, by decision makers, of more of the results currently available from models to supplement the predictions of the highest second-highest value (or a substitute as discussed in 9 above). Specifically, current modeling practice forces the calculation of concentration for a large number of time periods and at a large number of receptor points in order to identify the highest second-

highest values. This output provides valuable supplemental information to the decision maker on the frequency distribution of high values, as well as the geographic area of exposure and the relationships to meteorological patterns over different time intervals. This information, now largely ignored, should be used in decision making.

2

Review of Past Activity

2.1 INTRODUCTION

The first task undertaken by the Steering Committee of the AMS/EPA Cooperative Agreement on Air Quality Modeling involved a critical review of the process EPA used to develop the 1978 Guideline on Air Quality Modeling. The desire for such a document culminated in 1976 when EPA's Office of Air Quality, Planning and Standards (OAQPS) was directed to undertake its development. The Clean Air Act Amendments of 1977 further underlined this request mandating EPA to develop regulations specifying models for use in the prevention of significant deterioration (PSD) program. EPA developed a draft guideline document in late 1976 and eventually published a Guideline on Air Quality Models in April 1978.

To obtain an overview of the procedures leading to the Guideline, we studied the documents pertaining to the Specialists' Conference (February 1977), the Public Hearings (May/June 1977), and the National Modeling Conference (December 1977), all conducted to develop the Guideline. During the process of our review, EPA drafted a revised version of the Guideline (October 1980) and conducted public meetings (November 1980) on the new draft in Washington, Chicago, and Seattle. Each member of the Steering Committee spent numerous hours reviewing the more than 20,000 pages of formal documentation of this process. To be sure that significant ideas were not neglected, we retained Gordon Beals of Consolidated Edison of New york to prepare formal reports reviewing the comments received by EPA in 1977 and again in 1980. Dr. Beals' reports are excerpted freely throughout this chapter and are included as Appendixes A, B, C, and D.*

The process EPA used to arrive at the 1978 Guideline is described in the Guideline Preface as follows:

> "EPA undertook a series of steps that would lead to a widely reviewed guide on the use of air quality models. After initial opinions from EPA's Regional Offices were received, the Office of Air Quality Planning and Standards prepared a draft guide. This guide was submitted for critical review to a conference of specialists (Roberts, 1977). The individual conferees were widely recognized experts in the development and use of air quality models. Based on the judgments and suggestions of the conferees, the guide was revised and presented for public comment (Slater, 1977) at meetings in Atlanta, Chicago, Denver, New York and San Francisco. These meetings were attended by approximately 500 representatives of control agencies, industry, environmental groups and the scientific community. These attendees submitted extensive oral and written comments which were evaluated and considered in the preparation of this guide.
>
> "During the development of the guide, the Clean Air Act Amendments of 1977 were signed into law. These Amendments required EPA to conduct a Conference on Air Quality Modeling (Environmental Protection Agency, 1977a) with special attention given to models applicable to prevention of significant deterioration. That Conference was held in Washington, D.C. on December 14-15, 1977, with the 'Interim Guideline on Air Quality Models' (October, 1977) serving as the focal point. The Conference was attended by over

---

*The inclusion of these reports does not imply endorsement by the Steering Committee of all the information reported, nor does it suggest that the comments either made, or reported on, are valid.

> 300 individuals with similar backgrounds and interests to those who participated in the five previous meetings. A verbatim transcript of the proceedings was maintained and written comments were accepted."

2.2 SPECIALISTS' CONFERENCE

The EPA chose to focus comments on the first public draft of the Guideline through the mechanism of a "specialist" workshop. The workshop was conducted by Argonne National Laboratory in February 1977. Some 24 individuals, representing a wide range of the air pollution modeling community, were called together to review the draft guideline. It is informative to summarize the major considerations of that group.

2.2.1 Policy Issues

A number of policy issues were discussed in general sessions. Those relevant to the tasks discussed here include the following.

2.2.1.1 Guideline Tends to Standardize Air Quality Analysis

In recognizing that the Guideline tends to standardize models, the conferees noted that standardization would be acceptable for certain circumstances, if the particular models have been appropriately verified (or validated). The need for validation procedures was deemed a paramount requirement. Identifying concerns with "scientific" versus "regulatory" needs, the reported consensus on this issue is worth repeating. Namely (Roberts, 1977, p. 25),

> "...the conferees have reached consensus on this issue in that it is agreed that for certain standard types of situations, it would be desirable for the Guideline to identify an 'appropriate' or, possibly, 'approved' set of algorithms or particular computer codes, to be applied _if_ the particular model(s) have been properly verified under the conditions for which the standard application would be prescribed. The decision in this regard would be made in accord with an approved set of procedures for testing, validation, and possibly

> calibration. Secondly, significant flexibility should be provided to the user whereby, with appropriate professional counsel, other analytical methods could be employed where the standard situation was not encountered.

Thus the specialists clearly identified a condition on which the acceptability of a guideline depends. The guideline should be based on an accepted set of criteria for testing and evaluating models.

2.2.1.2 Specification of Computer Codes versus Approved Algorithms

The conference report (Roberts, 1977, p. 26) states

> "...the Guideline should, where appropriate, reference specific computer codes, although evaluation studies should be conducted on individual algorithms as well as on such complex air quality models. The working groups, however, had many reservations about the specific models listed in the Guidelines, due in part to the lack of sufficient verification experience, and in many cases could not give blanket approval to their use."

Again a key point comes through that model validation is held up as the missing component in the approval of models.

2.2.1.3 Application of Standard Computer Codes for "Screening" in the New Source Review Process

In view of the problems encountered in approving models, the concept of a conservative model (one that overestimates concentration values), used as a test or screening procedure, was generally supported. It was anticipated that much detailed and costly analysis could be eliminated if a simple screening model, sufficiently conservative, could be employed to approve sources. Cases that cannot be so approved would require a more refined analysis eliminating the conservatism of the screening model. This approach was adopted in the 1978 Guideline. Because of subsequent legislation and regulation, screening analysis difficulties have been encountered.

The 1977 Clean Air Act Amendments established numerical air quality increments for PSD that permitted applications. The use of screening techniques for the calculation of these increments may show increment consumption that would not occur with a more refined technique.

2.2.1.4 Evaluation and Calibration

Comments on this subject include (Roberts, 1977, p. 27):

> "Currently there is no systematic procedure being employed by the EPA for the evaluation of air quality models. Such a procedure would require criteria for acceptability and for inclusion in a guideline, procedures for carrying out a test program including specifications on data and on the range of condi- tions under which the model must be evaluated."

The issue of calibration procedures was also addressed and after specific subgroup deliberations was supported

> "....for evaluation of long-term average concentrations in urban, multi-source configurations."

2.2.1.5 Periodic Review and Upgrading of the Guideline

> "By recognizing the generally unsatisfactory status of past efforts in validation of air quality models and the need for a more extensive formalized procedure, the conferees wish to underscore the need for periodic review and upgrading of the Guideline."

This recommendation was taken up by the EPA, which is currently upgrading the modeling Guideline.

Other general issues raised at the specialists' conference were referred to working groups and did not finally evolve to any particular resolution.

2.2.2 Substantive Modeling Issues

The specialists' conference reported through eight working groups on substantive modeling issues. Specific recommendations were made for the following:

2.2.2.1 Multisource, Short-Term and Long-Term Nonreactive Pollutants

This subgroup suggested that packaged test cases should be developed including emissions data, meteorological data, and measured air quality data for testing models. They recommended that the Guideline clearly delineate limits of applicability of recommended models. Further, the subgroup suggested that EPA should "encourage continuing effort...for validating the models they are recommending." Moreover, the group suggested that "a formal procedure should be developed for approving new models and placing them in the guideline."

2.2.2.2 Single-Source, Short-Term and Long-Term Nonreactive Pollutants

Specific concerns raised included: questioning the validity of using Class A stability for tall stacks; the assumption of a horizontal wind adjustment factor in the CRSTER* model may not adequately depict σ_y ; 24-hour predictions from CRSTER may be too low; terrain adjustment must be decided on; atmospheric stability (or associated turbulence levels) should be a function of source height; some questions relating to calculating plume rise remain; and concern was expressed about applying algorithms beyond 50 km. This group suggested a screening procedure similar to the analysis in Guidelines for Air Quality Maintenance Planning and Analysis (Environmental Protection Agency, 1977b). Finally, the group made limited recommendations about statistical treatment of short-term averages.

2.2.2.3 CO, NO_2, and Photochemical Oxidant Pollutants

This working group recommended the identification of air quality models in terms of generic classes only. They proposed a two-level hierarchy of model usage-screening models for general estimation and potential problem identification and refined models that attempt to give the "best" prediction. The group also recommended screening "tools" or procedures for CO and NO_2 models.

---

*CRSTER refers to a specific air quality model described in Guideline on Air Quality Models (Environmental Protection Agency, 1978).

2.2.2.4 Long-Range Transport and Loss Mechanisms

A specific statement recommended cautions about application of long-range transport models to PSD calculations. This paragraph was adapted with little change in the final guideline.

2.2.2.5 Plume Dynamics under Special Conditions

Aerodynamic downwash was discussed with the suggestion that the "two-and-a-half times" rule should not be applied arbitrarily. Rather, consideration of relative heights of stacks and nearby structures or land features, dimensions of building and stacks (aspect ratio), emission characteristics (buoyancy), and stack configurations should all be considered. Sea breeze and like phenomena were considered to be a significant missing element from the Guideline. Interaction of the plume with an elevated inversion layer was seen as a problem area where models need improvement; however, specific recommendations could not be made.

2.2.2.6 Complex Terrain

Detailed discussion was provided on the issue. No particular existing model could be recommended under all circumstances. Thus a detailed process ranging from with a screening technique for the simplest cases to an actual data-gathering study for the most difficult situations was recommended. Recommendations were made to upgrade efforts to develop new models in this area.

2.2.2.7 Characterization of Turbulence

A number of detailed comments were made in this section that are essentially similar to the recommendations on the subject presented in the AMS Workshop on Stability Classification (Hanna *et al.*, 1977).

2.2.2.8 Validation and Calibration

In essence this group called for establishment of a specific continuing group within EPA charged with model validation and evaluation responsibilities. They also recommended that air quality models should meet certain standards of performance if they are to be endorsed in a Guideline for general use.

2.2.3 Conclusions of AMS Steering Committee on 1977 Specialists' Conference

The Steering Committee, based on a review of the specialist conference proceedings, offers the following general observations:

2.2.3.1 The specialists' conference was a useful mechanism to obtain substantive review of the draft guidelines. It should be retained in the process for future guideline development.*

2.2.3.2 Comments made at the specialists' conference were remarkably broad and perceptive. In general, they included all the points raised in the lengthy public hearing process.

2.2.3.3 There remain a number of technical concerns raised at the 1976 Specialist Conference, in the NCAQ Panel Report, and at the recent public meetings on the 1980 draft revised modeling guidelines, which have not been resolved. Examples include the following:

2.2.3.3.1 EPA models use different methodologies, particularly with regard to the final height to which a plume rises and the distance to that final height. There does not appear to be a strong scientific basis for using such different procedures. The UNAMAP4** model listing, however, corrects this deficiency.

2.2.3.3.2 Dispersion coefficients were originally developed from measurements averaged over 3-10 min. Models usually apply these coefficients to calculate hourly averages. The coefficients should be adjusted to account for these average differences.

---

*In November 1979 the National Commission on Air Quality convened a Dispersion Model Evaluation Panel. Although not conducted by EPA, it can provide input for the 1980 draft Guideline revision. Major recommendations from this Panel have recently been published (Fox and Fairobent, 1981). There is a strong parallel between recommendations made at each of these specialists' meetings.

**UNAMAP4 Is a computer tape of air quality models available from NTIS (5285 Port Royal Road, Springfield, VA 22161) provided by EPA to model users.

2.2.3.3.3 The general use of mixing depth information in EPA models should be improved, especially to allow a plume to penetrate partially into an inversion.

2.2.3.3.4 Guidance on the appropriate distance downwind that models can be applied correctly should be provided. Even though the 1978 EPA Guideline suggests that analyses be limited to less than 50 km from a source for other than Class I impacts, the scientific validity of such distances remains questionable.

2.2.3.3.5 Numerical SO_2 decay rates should be specified in appropriate models.

2.3 HEARINGS AND THE NATIONAL MODELING CONFERENCE (1977)

The record of the 1977 hearings conducted in New York City, Atlanta, Denver, San Francisco, and Chicago, including formally submitted material, is very lengthy, on the order of 10,000 pages. A detailed summary prepared by Gordon Beals is included in Appendixes A, B, and C.

Among the questions raised at these meetings are the following:

2.3.1 What is the appropriate degree of specificity and of flexibility that the Guideline should adopt? This issue is at the heart of controversy over the review and selection process of new models discussed in Chapter 4.

2.3.2 Is consistency of modeling between regions needed? We have suggested that EPA may be misconstruing concerns about the need for consistency. This point, including recommendations on its resolution, is discussed more fully in Chapter 4.

2.3.3 What is the appropriate use of air quality data? Air quality data are considered in two different contexts. Air quality data are needed for model evaluations. Data availability and quality are major factors in the success of model evaluation programs. On the other hand, a number of commenters recommended increased use of air quality data in the PSD permitting process. These commenters did not seem to recognize the fact that the PSD permit is, by statute, based on a model calculation. Any discussion on the merits of data versus modeling is beyond the scope of this report.

2.3.4 How shall models be evaluated/validated? Again this issue has received major emphasis by the Steering Committee and is discussed in Chapter 4.

2.3.5 What constitutes representative meteorological data? This question cannot be answered easily. It is discussed in Chapter 5.

2.3.6 How should models deal with variable emission rates and operating loads? This issue is expected to receive considerable attention at the 1981 National Modeling Conference. It relates more to the regulatory decisions made using model results than to the modeling itself and therefore was not addressed.

2.3.7 What are the most appropriate dispersion coefficients for various models? Steering Committee comments on this issue are embodied in Chapter 3 and in the Appendixes.

2.4 PUBLIC HEARINGS ON THE 1980 DRAFT GUIDELINE

In October 1980, EPA issued a revised draft modeling guideline. This draft was subjected to public meetings and a public comment period. Information meetings were held in Washington, D.C., Chicago, and Seattle. Extensive formal documentation of comments and concerns was submitted. The Steering Committee again asked Gordon Beals to review the public record. His report appears as Appendix D.

Comments were solicited by EPA in eight specific areas, namely:

1. Specific changes to EPA recommended models;
2. Screening and refined techniques for complex terrain;
3. Requirements for emissions, meteorology, and air quality data bases, including the determination of background;
4. Selection of receptor sites;
5. Criteria for the case-by-case acceptance of nonguideline models;

6. Use of measured data in lieu of model estimates for establishing emission limits;

7. The appropriate number and placement of monitors to support a model evaluation;

8. Performance measures and standards for models.*

In the following subsections the public comments responsive to the most significant of these areas are outlined.

2.4.1 Specific recommendations for changes in EPA Guidelines were made in numerous (over 200) comments. They include comments on plume rise, treatment of complex terrain, dispersion parameterization coefficients, inconsistencies between urban and rural formulations, downwind limitations, and the adequacy of the Gaussian concept. Detailed comments are included in Appendix D.

2.4.2 What models should be recommended for screening and refined analysis in complex terrain? While there seemed to be general dissatisfaction with the use of the Valley model and especially with the meteorological inputs it requires as a screening technique, there was no obvious consensus about what should be used in its place. Research is needed to specify a replacement. Use of the the Valley model is anticipated to be particularly controversial for calculation of PSD increments in complex terrain situations. Since the model is recognized to be conservative, the likelihood of premature and inaccurate increment consumption exists. The use of such conservative screening techniques, therefore, does not seem acceptable for PSD calculations in which multiple sources are anticipated for the same area.

2.4.3 What emissions, meteorological, and air quality data should be required for model applications? The 1980 draft Guideline recommends a requirement for five years of meteorological data from a National Weather Service (NWS) site or one year of site-specific data. Considerable dissatisfaction was expressed concerning EPA's strict requirements to determine the representativeness of the data. Many commenters expressed concern about the validity and expense of collecting such data.

---

*We believe that EPA should pursue the recommendations that we have made with regard to the performance evaluation and peer review of models, including both new techniques and those already contained in the guideline, because this approach would go a long way toward providing suitable response to specific technical issues.

2.4.4 What criteria shall be used for accepting nonguideline models for regulatory applications? There was considerable concern that EPA's requirements suggested in the 1980 draft guideline for the use of a nonguideline model are not conducive to the acceptance of new, improved models for general use. Requiring equivalency within 2 percent of the numerical value of a model basically requires that alternative models be nearly identical to the guideline models. The tacit assumption that guideline models are norms or standards against which all other techniques should be tested was attacked in the context that these models had not themselves gone through thorough performance tests. The alternative route, of having to demonstrate on a case-by-case basis and through the use of extensive site-specific data that a nonguideline model is superior, was considered to be a major obstacle to the actual adoption of new models.

2.4.5 What are appropriate performance measures and standards for models? While there was general support for the concept of model standards, no one put forward any specific recommendations. A number of performance measures, however, were suggested.*

---

*All the suggestions appear to be covered by the techniques recommended at the AMS/Woods Hole workshop (Fox, 1981).

3

State of the Science of Air Quality Modeling

This chapter is divided into three parts. The first part briefly reviews the scientific basis for modeling and the known performance characteristics of existing models. The second part summarizes six review papers on the state of the science of air quality modeling. The third gives an overview of research needed in several specific areas related to air quality modeling.

3.1 SCIENTIFIC BASIS FOR MODELING AND KNOWN PERFORMANCE CHARACTERISTICS

The science of air quality modeling provides a link between emissions and observed ground-level concentrations of pollutants. As meteorological conditions such as wind speed and stability vary over typical ranges, the ground-level concentration is known to vary over several orders of magnitude at a given location for constant emissions. Using known physical concepts such as the principles of mass continuity and the equations of motion, scientists have developed a variety of air quality models that can satisfactorily simulate observed concentration patterns. More importantly, the models can provide estimates of pollutant concentrations in areas where no monitors exist. Thus the models have a solid scientific basis and can be used for decision making for hypothetical situations.

Given that the models are scientifically valid, the next questions are: How well do they perform when compared with data? How close do they get you to the real answer? First, it must be stressed that the model is not the only factor contributing to differences between predicted and observed concentrations. There are also contributions due to errors in model input parameters and errors due to turbulent variations in observed concentrations. Pasquill (1974) summarizes the expected errors in models applied to a variety of situations. For research-grade observations of diffusion from near-surface sources over flat terrain, the expected error is 10 to 20 percent. In most real-world model applications to elevated plumes with only a few supporting meteorological data, Pasquill estimates that the expected error (for 1-h averages) is about a factor of 2. Long-term means in this situation are accurate within about 20 percent.

Recent detailed experiments in St. Louis (Ruff, 1980) and Illinois (Bowne, 1981) suggest that standard (i.e., EPA) models are actually less accurate for 1-h averages than Pasquill suggested in 1974. Correlation coefficients between observed and predicted concentrations at fixed stations are low, and errors greater than a factor of 2 are common. We believe that this error is mainly due to a failure to estimate the wind direction adequately. That is, the plume centerline observed and estimated concentrations are within the expected error suggested by Pasquill, but the location of the centerline might be off by 20°. Some of the performance measures in Chapter 5 look only at the magnitude of maximum concentration predictions and observations rather than at their location.

3.2 SUMMARIES OF REVIEW PAPERS ON THE STATE OF THE SCIENCE

The AMS Steering Committee sponsored several papers on the state of the science of air quality modeling. Some of these (Bass, 1980; Bowne, 1980; Hosker, 1980b; Latimer *et al.*, 1980) were presented at the 1980 AMS New Orleans Second Joint Conference on Applications of Air Pollution Meteorology and published in the proceedings volume. Other papers, by Smith and by Randerson, were not published externally but were used in the preparation of recommendations in this report. All of these papers are included in Appendixes E through J. In this chapter a brief abstract of each paper is given, along with one or two specific recommendations.

3.2.1 A. Bass, Modeling Long-Range Transport and Diffusion (Appendix E)

Several long-range transport and diffusion models intended for regulatory applications are reviewed, including those used in Europe for SO_2 modeling. Most models follow puffs along trajectories calculated from observed wind fields. Simple methods are used for estimating diffusion and dry and wet deposition. Bass concludes that trajectory-type models are better tools for regulatory decision making on long-transport scales than are the conventional short-range, straight-line Gaussian plume models. He also stresses the importance of gathering new data, since the available data base for evaluating long-range models is meager.

3.2.2 N. Bowne, Validation and Performance Criteria for Air Quality Models (Appendix F)

Performance measures for evaluating air quality models are summarized. These include calculations of the mean and standard deviation of the set of predicted minus observed concentrations at a point, the time and space correlation between predicted and observed concentrations, and the analysis of predicted and observed frequency distributions of concentrations. Measures for evaluating the ability of a model to predict the magnitude and location of maximum concentrations are also given. Much of what Bowne discusses was used for the Woods Hole Performance Evaluation Workshop, which is described more fully in Chapter 4.

3.2.3 R. Hosker, Dispersion in the Vicinity of Buildings (Appendix G)

A review is given of a large number of wind tunnel and field experiments on the subject of dispersion in the vicinity of buildings. This research area is divided into three subgroups: source upwind of a building, source on or near a building, and source downwind of a building. Recent studies show that flow around buildings is more complicated than previously thought; for example, persistent vortices behind a building can strongly influence dispersion rates at distances of up to 30 times the building height. To develop schemes to account for dispersion around buildings, it is recommended that the EPA study Hosker's conclusions and his comprehensive list of references.

3.2.4 D. Latimer, R. Bergstrom, C. Johnson, and J. Killus, Modeling Visibility (Appendix H)

Work done by the authors in preparing a workbook for the EPA to assess visibility impairment in Class I areas is reviewed. Class I areas include national parks and other wilderness areas that require special protection. The authors state that visibility models are still in their infancy but that much research is currently under way in the development of model components. Transport and diffusion to downwind distances of 100 or 200 km must be calculated, as well as chemical transformations, aerosol growth, radiative properties of the plume, and response of the human eye. The problem of transport and diffusion is identified as the critical problem in this research area. It is recommended that it may be possible to protect visibility through existing NSPS and PSD regulations or through surrogate measures such as plume NO_2 and fine particulate concentrations.

3.2.5 D. Randerson, Wet and Dry Removal, Resuspension, and Atmospheric Chemistry of SO_x and NO_x (Appendix I)

The author presents an overview of the subject of removal processes, including wet and dry removal, resuspension, and chemical removal. Values for deposition speeds, wet removal rates, and chemical reaction rates are given for many common pollutants. It is clear that these parameters need to be refined by improved field experiments and applications of theoretical models. Long-range transport and chemical transformations are addressed in relation to visibility degradation and acid precipitation. A long list of recommendations is given, most of which emphasize the need for more data in the wet and dry removal area and more data and better theories in the chemistry area. Errors associated with removal parameters need to be decreased so that the parameters can be more confidently chosen for use in regulatory models.

3.2.6 M. Smith, Transport and Diffusion Modeling and Its Application--1980 (Appendix J)

This paper reviews the history of transport and diffusion modeling and problems with applying models in the present regulatory environment. Smith believes that the necessary field experiments and model evaluations are lacking for the set of EPA models being applied today and that many persons applying the models do not understand their range of applicability. He makes the strong statement that models are frequently being misused, largely based on the actions of overzealous legislators and regulators. Both the advantages and limitations of models need to be recognized before misuse of models discourages their proper application.

3.3 RESEARCH NEEDED TO IMPROVE AIR QUALITY MODELS

In this section research needs related to air quality modeling are outlined, with special emphasis on EPA needs. Sixteen specific issues are addressed, including new topics such as long--range transport and visibility models. Research needs of the five different types of diffusion models will be discussed, including the statistical model, the similarity model, the diffusion equation, the Langevin equation, and the Gaussian plume model. The order of listing of these issues and models is determined by scientific reasoning rather than by research priorities. In addition, research needs for diffusion modeling in special circumstances are discussed.

First we are concerned that there is currently much more application than basic research under way in air quality modeling. To advance the field, more basic research must be supported by both government and industry. Models should be based on an improved theoretical basis and scientific evaluation. Unfortunately, the field has tended to go ahead with the risky business of extrapolating limited existing knowledge into new problem areas where the physics is uncertain and the data not available. New problems are developing at a faster rate than are the basic research programs to consider them (e.g., visibility, complex terrain, deposition). By basic research programs, we mean not only the collection of data but also the thoughtful analysis of these data and comparison of data with the predictions of theoretical models. More research work should be supported on new concepts at the frontier of our science.

In addition, we wish to note that there are many diffusion data sets in existence that have not been thoroughly analyzed. Before embarking on expensive new field programs, these old data sets should first be searched and reanalyzed. For example, there are the St. Louis Regional Air Pollution Study (RAPS) data set, the Pennsylvania-LAPPES data set, and the Tennessee-Cumberland Steam Plant data set, just to mention a few. Many have thought that the 1958 Prairie Grass data set had been completely analyzed, but Briggs and McDonald (1978) recently successfully reanalyzed these data using the mixing height z_i as a scaling parameter.

3.3.1 Similarity Models of Diffusion

In similarity models of diffusion, diffusion parameters such as the standard deviation of concentration distribution, σ_z, and the mean height of the pollutant plume, z, are functions of basic meteorological parameters such as the mixing depth, z_i, the friction velocity, $u*$, the Monin-Obukhov length, L, and the roughness length, z_0. Boundary-layer experiments during the past decade have shown the usefulness of the similarity approach. In convective conditions the convective velocity scale, $w_* \equiv gH_0z_i/T_0$ has been found to be a useful scaling parameter. The parameters g, h_0, z_i, and T_0 are the acceleration of gravity, the surface heat flux, the mixing depth, and the absolute temperature, respectively. Briggs and McDonald (1978) have reanalyzed the Prairie Grass diffusion data from the 1950's and proved that the daytime data are better ordered when z_i and $w*$ are used as the scaling length and velocity. Several researchers have recently shown that the parameters $w*$, $u*$, L, z, z_i, and z_0 can be used to make precise estimates of stability class as input for diffusion estimates. The parameters u_*, i, z, and z_0 are the friction velocity, the Monin-Obukhov length, the height, and the roughness length.

Existing diffusion data sets should be reanalyzed to refine constants appearing in the similarity formulas and determine the expected accuracy of the formulas. Furthermore, whenever new diffusion experiments are conducted, it is important to take measurements of the parameters w_*, u_*, L, z_i, and z_0. This recommendation also came out in two workshops (Hanna et al., 1977; Hoffnagle et al., 1981).

3.3.2 Improvement of Gaussian Plume Model

The Gaussian plume model is expected to continue as a major tool in air quality analyses. It can be improved by recognizing that the initial diffusion of a large and buoyant plume (e.g., a typical power-plant stack) may be dominated by buoyancy rather than by ambient turbulence out to distances of several kilometers. Furthermore, methods of estimating σ_y and σ_z can be improved by making them functions of the standard deviations of wind-direction fluctuations σ_θ and σ_ϕ , or the boundary-layer parameters u_*, w_* z_i ,z_0,and L mentioned above in the similarity section. This recommendation was made by Hanna et al. (1977) and has been partially implemented by Irwin (1979). British models (Clarke, 1979) currently use the assumption that σ_y and σ_z are continuous functions of roughness z_0. Methods for estimating the effects of averaging and sampling times and release height on σ_y and σ_z need to be developed. More needs to be known about diffusion at downwind distances greater than 10 km, and new experiments and theory in this area should be supported. In particular, the influence of wind shears on diffusion at intermediate and long distances must be determined (Pasquill, 1976). See Smith's review article (Appendix J) for further discussion of the advantages and disadvantages of the Gaussian plume model.

3.3.3 Statistical Models of Diffusion

Computers are now fast enough to run Monte Carlo statistical models of diffusion, at least for special circumstances. In these models, the trajectories of thousands of particles are followed under the assumption that local particle speed is determined by the mean wind speed, a turbulent fluctuation that is correlated with the turbulent fluctuation at a previous time, and a random turbulent fluctuation. Calculation time steps are short because they must be less than the Lagrangian time scale, T_L. The model by Reid (1979) is an example of a modern statistical model. Research should be conducted to acquire the data and parameters necessary to run statistical models of the past. For example, three-dimensional distributions of mean wind, turbulent energy, and Lagrangian time scale are needed. Turbulence parameters for convective conditions can be obtained from the Minnesota experiment (Kaimal et al., 1977), but more field experiments are needed for neutral and stable conditions. Mean wind can be estimated from a three-dimensional mesoscale model or measured by a monitoring network. Some information is available on Lagrangian time scales (Pasquill, 1974; Hanna, 1981), but much more theoretical work and experiments are necessary. The statistical models should give us useful information on the internal structure of plumes and puffs, as well as on the variation of σ_y and σ_z with distance.

3.3.4 The Diffusion Equation

The diffusion equation, basically an expression of the conservation of mass for a pollutant, is a linear equation easily programmed for solution on a high-speed computer and therefore often used in applications. The modeler uses a multidimensional grid and needs initial and boundary conditions for pollutant concentrations. Also needed are source information (emissions and configurations), mean wind speed, and diffusivity coefficients (most generally a three-dimensional, second-order tensor; Calder, 1949) as a function of grid position. Current applied research on the diffusion equation centers on numerical techniques for reducing the formation of errors in the solution. Several types of finite-difference schemes are available (Egan and Mahoney, 1972), as well as other schemes in which moments or spectra of the distribution are used (Fox and Orszag, 1973; Prahm and Christensen, 1977). Investigators are continuing the work of Sherman (1978), who devised methods by which an observed mean wind field can be forced to be nondivergent. However, with the exception of work on second-order closure approximations of diffusion (Donaldson, 1973), little research is under way on the specification of the diffusivity field. Similarity theory and field studies in which fluxes of pollutants are measured in three dimensions in the full depth of the atmospheric boundary layer should be used to develop improved formulas for K. Research on methods of combining mesoscale meteorological models and diffusion models should be expanded.

3.3.5 New Diffusion Theories

The four basic diffusion models (statistical, similarity, Gaussian, K) have recently been joined by a new model, the Langevin model. This partial differential equation for modeling turbulent velocity fluctuations assumes that there are both correlated and random components of the fluctuation. The model has been applied in the fluid-dynamics literature for many years, but Gifford (1981) has demonstrated its utility in the atmosphere. With the Langevin model, observed σ_y curves can be simulated over several orders of magnitude of travel time. This model needs to be further developed and tested with field data. In addition, there may be techniques developed in other disciplines that are ripe for application to diffusion.

3.3.6 Diffusion and Wind Flow around Generic Building Shapes

Unless the pollutant is released from a tall stack with a height of about 2.5 times the neighboring building height, there is likely to be an interaction of the plume with the disturbed air flow around the building (Hosker, 1980b; Appendix G). There are data available on airflow and diffusion around specific buildings but only limited data for generic building shapes. By generic, we mean typical general shapes such as cubes or cylinders, turned at a wide range of angles to the mean flow, with a variety of typical emission source locations. These data could be collected in wind tunnels, water channels, and field observation programs. Concentrations in the wake and cavity regions should also be measured and basic theoretical similarity formulas developed.

3.3.7 Plume Rise

Briggs (1975) has outlined techniques useful for estimating plume trajectories and final plume rise in calm and windy stable conditions, and scores of field studies have verified his formulas. But there are still several problems in plume rise that require more research. For example, final plume rise in neutral and unstable conditions is not well known. Penetration of elevated inversions by buoyant plumes needs further work. Another problem is plume rise from multiple sources. At many locations the potential exists for nearby plumes to intercept each other and alter the plume trajectory. Most cooling towers and many power plant stacks occur in groups of two or more.

3.3.8 Maximum Ground-Level Concentration

Some EPA regulations are based on maximum ground-level concentrations for 1-h or 3-h averaging times. For short stacks (stack height less than about 100 m), highest concentrations occur in high-wind, neutral conditions, or morning fumigation periods, which are easily treated using the Gaussian plume model. Recent studies of diffusion from tall stacks show that highest concentrations occur in light-wind, very unstable conditions (e.g., Weil, 1979; Trout, 1979). In these conditions, the usual Gaussian model is no longer valid, since the plume is being brought to the ground in broad downdraft regions between thermals. Instead, modelers have had some success using convective similarity theory (Venkatram, 1980a; 1980b), theoretically finding that the plume maximum concentration descends with a speed of 0.4 w_* (Lamb, 1979). Boundary-layer scientists are developing an understanding of the intermittent convective boundary layer, and diffusion modelers should watch this work carefully to determine applicable results. It is hoped that field studies of diffusion from tall stacks in convective conditions will lead to refinement of the models with regard to both magnitude and location of maximum concentration.

3.3.9 Diffusion in Shoreline Environments

Many power plants and industries are on the shores of lakes or oceans. Because of the reduced roughness and the presence of waves, boundary-layer airflow and diffusion over water departs significantly from what we are familiar with over land. Possibly of greater importance are the sharp contrasts that occur at the shoreline in spring and fall when there is a great temperature difference between the land and water surfaces. For example, with onshore flow in the spring, the plume from a tall stack is emitted into stable air but usually intercepts a rising unstable layer, and fumigation to the ground occurs (Lyons, 1975). Limited field experiments have been conducted in this environment, and a few crude models have been developed, but there clearly is a need for more work. Since the environment is nonhomogeneous horizontally, the diffusion equation or a Monte Carlo statistical technique is needed to estimate detailed distributions of pollutant concentrations. It may be necessary to combine a diffusion model with a mesoscale meteorological model such as Pielke's (1974) model for the Florida sea breeze. These complex models should be used to develop a set of simple empirical models for practical use.

3.3.10 Diffusion in Complex Terrain

Many industrial facilities are being built in complex terrain, where meteorological conditions are nonstationary and nonhomogeneous. Complex terrain affects transport and diffusion in several ways. The air flows over and around hillsides; sometimes under sufficiently stable conditions the air flow is blocked by the hillside. The terrain can channel local winds so that wind directions can be markedly different from geostrophic wind direction. Turbulence levels for a given stability and roughness are greater over complex terrain than over a flat surface. Surface heating and cooling leads to upslope winds during the day and downslope winds during the night. Orgill (1981) has recently reviewed the state of knowledge in the field, adding to the review of these effects by Egan (1975). Unfortunately, few good field experiments have been performed to confirm these concepts. Drainage winds in the Geysers area of California (vertical relief 1000 m) are currently being studied by the Department of Energy (DOE) (Dickerson and Gudicksen, 1980), and diffusion around an isolated butte in Idaho (vertical relief 100 m) is being studied by EPA. Climatological research on an isolated 1000 m mountain has recently been initiated by the Forest Service (McCutcheon *et al.*, 1981). Careful analyses of these data should permit models to be improved. But there is still much to be done in studying a wide variety of hill and valley shapes for different types of sources.

3.3.11 Diffusion in Urban Areas

There was much activity in the area of urban diffusion modeling in the period 1965-1975. A detailed urban diffusion experiment, the RAPS program, took place in St. Louis during 1976-1977, but only small parts of the data have been analyzed so far (e.g., Ruff, 1980). Despite the proliferation of urban diffusion models and the large amount of data collected, there is still much to be done. For example, σ_y and σ_z curves for urban areas should be refined. Current curves are based on an earlier experiment by McElroy and Pooler (1968), and there is much room for improvement. Furthermore, urban diffusion models have difficulty simulating hourly SO_2 concentration variations at a single receptor when there are significant point sources in the area. There is the problem of increased concentrations in urban canyons (Johnson *et al.*, 1976), which is currently treated by a simple empirical model that needs to be tested further at many other locations. Multiple inversion layers are often present over an urban area at night, and models need to account for the trapping of pollutants by these layers. In other words, despite the funds expended on urban diffusion modeling in the past, urban diffusion models still have numerous problems. For this reason, AMS and EPA are planning a workshop conference in 1982.

3.3.12 Long-Range Transport and Diffusion

Bass' review (Appendix E) provides a rather complete look at modeling--particularly from the regulatory perspective. The NCAQ (Fox and Fairobent, 1981) also addressed this subject, stating that the significance of both SO_2 conversion and oxidant formation separately and in combination should not be neglected in air quality management. They called for increased use of chemical kinetics and deposition parameters to make areal estimates of pollution impact. A central question that remains is whether a specific source-receptor relationship can be established with sufficient confidence to determine emission limitations. Johnson (1980) recently reviewed the literature and concluded that research efforts should continue in the following areas: (i) incorporating complex terrain effects on transport wind and deposition; (ii) improving vertical resolution, boundary-layer processes, and vertical diffusion; (iii) improving deposition formulations; and (iv) incorporating satellite-derived data. Extensive research work is planned as a result of current public concerns over acid precipitation (Interagency Task Force on Acid Precipitation, 1981). The Steering Committee strongly endorses these efforts but not without a word of caution. While the particular regulatory context changes from year to year, scientific problems do not change quite so rapidly. Research takes time as well as support and people. In the area of long-range transport, air quality modeling is beginning to come close to the smallest scales of the meteorological mesoscale research efforts, perhaps the most active in the atmospheric sciences. DOE, EPA, and a host of other agencies are involved. Coordination to establish a truly meaningful research effort will be a major job. For these reasons, the Steering Committee is planning a workshop to deal with the meteorological aspects of long-range transport and diffusion.

3.3.13 Visibility Models

A component of recently developed PSD regulations is the requirement that visibility be protected in national parks and wilderness areas. Latimer et al. (1980; Appendix H) reviewed the state of the art in visibility modeling and noted that it is still in its infancy and that it could benefit greatly from new research programs. Present regulations only address visibility impairment by plumes from single, identifiable sources (Federal Register, 1980, pp. 34762-34779, May 22, 1980). Additional research, data collection, and instrumentation will be needed to provide reliable answers to regional-scale visibility problems.

A visibility model is first based on a diffusion model in which chemical mechanisms are included to account for chemical reactions, gas-to-particle conversions, and particle growth. An atmospheric optics submodel is applied to the output of the diffusion model, accounting for light scattering and absorption, background effects, and the location of the plume observer. Additional theoretical and experimental work is needed in each of these areas.

3.3.14 Chemical/Reactions

In all situations where the primary pollutants are chemically changed or where new secondary pollutants are formed through chemical reactions, the diffusion models must include chemistry. Urban smog chemistry has been studied extensively, and an example of a recommended set of reactions is reported by Falls and Seinfeld (1978). Sulfur plume chemistry is a big part of the DOE-EPA MAP3S program and the EPRI-sponsored SURE program, both concerned with the atmospheric transport of sulfur in the eastern United States (see Appendix I). Long-range transport and formation of ozone is the subject of recent EPA field programs in the Northeast. Many of the chemical rate constants have been estimated from laboratory chamber experiments, and the question often comes up whether these rate constants are valid in the open atmosphere, in the presence of other catalysts and boundary effects. A major problem is what happens to the rate constants at night.

3.3.15 Wet and Dry Removal Processes

PSD regulations applied to wilderness areas and national parks require an estimation of the impact that an air-pollution source is likely to have on the ecosystem of the area. This impact cannot be properly assessed without a model that provides for the deposition of pollutants from the air to the ground, to surface waters, and to leaves of vegetation. In addition, current concerns over acid rain and the long distance transport of pollution suggest a need for increased understanding of both wet removal (precipitation and cloud) and dry deposition to surfaces.

A recent survey (Hosker, 1980a) illustrates the current understanding of the physics of deposition. Simple parameter-laden empirical formulas are available and are capable of expressing the deposition mechanisms. The appropriate parameters for a few pollutants (McMahon and Denison, 1979) have been measured. However, reviewers such as Sehmel (1980; 1981) and Slinn _et al._ (1978) point out that there is a wide range of uncertainties involved in these estimates and a lack of knowledge of parameters for deposition at night, deposition over water, and removal of toxic metals and trace organics. A recent workshop in this field supports this point (Hicks _et al._, 1980). It appears that an understanding of the fundamental turbulent interactions within the plant canopy (Thom, 1975; Wilson and Shaw, 1977; Lewellen and Sheng, 1980) will be needed to progress in this area.

3.3.16 Long-Range Transport and Diffusion

Because of concerns over the long-range transport and diffusion of sulfates, nitrates, radioactive material, and other compounds, interest in this speciality has experienced a rapid growth during the past decade. Several models have been developed, most of them extensions of the standard Gaussian formula (Bass, 1980; Appendix E). In general, the physical assumptions in these models are quite crude. For example, Heffter et al. (1975) assume that $_y$(m) = 0.5t (sec) for all applications. Transport is given by interpolations within an observed wind monitoring network. Serially released puffs are assumed to grow as their centers of mass are transported by the wind field. A major research problem is that plume 's are used for estimating plume diffusion. To properly define puff 's, dry and wet removal rates, and wind velocity interpolation schemes, there is no substitute for long-range tracer experiments under a variety of meteorological conditions. With regard to defining the wind field, the proper "representative" number of wind-monitoring stations needs to be defined.

4

The Evaluation of Air Quality Models for Regulatory Applications

4.1 INTRODUCTION

One of the many uses of air quality modeling is for regulatory applications. There are a variety of expressed needs associated with these applications. Two needs especially appear to be in conflict at present, and this conflict has been the focus for much of the activities of the steering group and for much discussion at various modeling conferences and technical meetings. On one hand, regulatory applications require model predictions to be as correct as possible. Because of the magnitude of the economic and environmental implications of decisions made on the basis of such predictions, the model should represent as close a simulation of how nature behaves as is practical. On the other hand, the regulatory process is simplified if there is consistency between models in use. Consistency leads the regulating agency to be inflexible with regard to recommending new techniques.

For regulatory application, a model must be consistent and correct. Consistency and inflexibility carried to an extreme imply the use of one single model for all applications for all time. On the other hand, correct modeling requires continuous alteration and changes as new findings are made and improved methods are formulated. In selecting a modeling approach for regulatory application one must somehow draw a balance between these two extremes. It is important that the best new approaches are able to replace less correct techniques, but it is just as important that equally valid models not provide estimates so widely divergent that decisions on a specific emission limitation cannot be made effectively.

By issuing the 1978 Modeling Guideline, EPA took a step toward consistency. Since publication of the Guideline, most of EPA's activities with regard to controversial modeling quesions have continued the trend toward consistency. However, the Clean Air Act requirements for a Modeling Conference every three years implies that guidelines should be upgraded on a regular basis. The scientific community recommended that the selection of models for inclusion in the guideline be based on a test of model performance and sufficient indication that the model perform up to some standard. Precisely what these standards should be is a question of considerable complexity.

It is therefore necessary to develop a procedure to accomplish an objective and scientifically meaningful evaluation of model performance. As well as being necessary from a scientific viewpoint, the public interest is served best by models that will neither systematically overestimate nor underestimate the impacts.

For many applications today we simply do not have sufficient information to know to what extent a model overpredicts or underpredicts. Therefore, we need to improve our ability to simulate the actual conditions and to be able to improve the confidence in our results. Control programs will improve as a result of greater confidence being placed in model predictions.

The nature of environmental protection is an evolutionary one. As scientists our goal is straightforward, namely, to understand the relationship between emissions and air quality, to document the health and welfare consequences of air quality deterioration, and to develop methods for ameliorating such deterioration. The public, through the political process, decides how much deterioration is acceptable and how it should be prevented.

Not only is it important to improve models, but it is equally important to improve the manner in which models are applied. As we have noted earlier, regulatory applications of models often do not utilize the most appropriate information provided by a model. Some reform is necessary in the manner in which decision makers use model outputs.

It is apparent that a balance must be drawn between consistency and correctness. Models must be upgraded and improved, but some degree of consistency in their application must be maintained. Issuing a guideline that establishes standards for routine modeling applications has provided a degree of necessary consistency. Given a guideline, how should one model be selected over another for inclusion in the guideline and for use in a particular regulatory activity? This important question is addressed in the remainder of this chapter, where we discuss the methods of soliciting new models, of developing a procedure for evaluating one model against another, of specifying the principles on which one model can be selected over another, and finally of reviewing some of EPA's approaches with regard to particular decisions on models and their application.

4.2 SOLICITATION OF NEW MODELS

The 1978 Modeling Guideline was criticized for promulgating only models that were developed by, or for, EPA. As a result of this concern and with the encouragement of the AMS, EPA conducted an open solicitation for new models (Federal Register, Vol. 45, No. 61, March 27, 1980, pp. 20157-20158). EPA placed some stringent requirements on model submission, which basically required that a submitted model be in a finished form (i.e., fully documented, computer compatible with EPA, test cases).

The submission of a scientific publication containing, for example, a new concept and an indication of its validity would not be adequate. The need to consider only "finished" models is understandable from a regulatory perspective. However, this requirement means that new ideas would emerge, in practice only, through EPA's Office of Research and Development (ORD) or through operationally oriented model developers such as contract research and consulting companies or government laboratories. The Steering Committee is concerned that these constraints will slow the emergence of new technologies.

Perhaps some formal mechanism for the consideration of new ideas and research findings in ORD would be appropriate.

As a result of the solicitation, between March 1980 and April 1981, 28 new models were submitted to EPA for consideration. At the October 1980 public meetings EPA proposed to classify the models received as either:

1. Recommended for use in specific applications;
2. Allowed if shown to be "equivalent" to an existing recommended model ("equivalence" was proposed to mean that computed values would be within 2 percent of those computed with a recommended model);
3. Allowable if a case-by-case demonstration based on verification with data and in accordance to detailed requirements in Section 6 of the proposed guideline shows it to be more applicable in a given situation;
4. No recommendation; or
5. Rejected.

While we support EPA in their open solicitation for models, we are concerned about what will happen subsequently. The EPA proposed actions are keyed in large part as to how a submitted model compares with an existing EPA recommended model. Since the EPA models have not themselves been subjected to sufficiently comprehensive performance evaluations to be established as standards, this rationale does not focus on scientific merit. The equivalency test proposed would reject all models that are not basically identical to the "recommended" models. Therefore, advances in science would not emerge from this process. The requirements for case-by-case demonstrations are very stringent and require someone with an alternative model to gather substantial quantities of new site-specific data as a basis for the demonstration. In this process the costs involved may be a disincentive to the developer. At best, a new model might be approved for the specific site. We note that no procedure is identified for the upgrading of or the replacement of a recommended model in time. Our concern is that the determination needs to be based on scientific merit and performance measures rather than consistency. EPA needs to show more willingness to accept innovations in modeling techniques that have promise in providing more accurate predictions of air quality values.

4.3 PERFORMANCE EVALUATION

In evaluating one model against another, performance must be a major consideration. There has been significant activity in the past few years directed toward the evaluation of air quality models. Notable among these efforts has been work by EPA (Hayes, 1979; Hillyer et al., 1979) and by the Electric Power Research Institute (EPRI) (Hilst, 1978) to develop procedures for evaluating models. Issues relating to performance evaluation were the subject of a recent workshop that the Steering Committee conducted at Woods Hole in September 1980. A report (Fox, 1981) from that Workshop appears as Appendix K. While the Woods Hole Workshop report focuses heavily on statistical evaluation procedures, it also recognizes the need for scientific evaluation, stating, "A scientific evaluation involving examination of cause and effect relationships...-should be included." In this context, we distinguish between a purely statistical and quantitative comparison between data and model estimates and the qualitative acceptance of general physical principles (e.g., conservation of mass), which we term scientific.

4.3.1 Statistical Evaluation

The concept of evaluating model performance, by comparing a sequence (a series of numbers over time) of model predictions against measured values, is straightforward. The techniques and procedures are well founded in time series analysis (Box and Jenkins, 1976; Londergan, 1981). The essence of the evaluation is based on a set of model predictions, $c_p(x,t)$ and a set of observations, $c_o(x,t)$ of ambient air quality.

A quantitative comparison between c_p and c_o is based upon a statistical examination of the differences, defined as

$$d = C_o - C_p .$$

In particular, the average (bias),

$$\underline{\bar{d}} = \frac{1}{N} \Sigma \underline{d} ,$$

the variance (noise)

$$S_d^2 = \frac{1}{N - 1} \Sigma (d - d)^2 ,$$

and the gross variability (gross error) expressed as, either

$$|\bar{d}| = \frac{1}{N} \Sigma |d|$$

or

$$\mathrm{MSE}\ d = \frac{1}{N} \Sigma\ d^2 \qquad \text{(mean-square error)}$$

are suggested to measure the magnitude of difference between the model predictions and observations.

The magnitude of correlation between them is measured by the correlation coefficient,

$$\gamma_{op} = \frac{\Sigma\ (c_o - \bar{c}_o)\ (c_p - \bar{c}_p)}{\left[\Sigma\ (c_o - \bar{c}_o)^2\ \Sigma\ (c_p - \bar{c}_p)^2\right]^{\frac{1}{2}}},$$

which can be calculated for space, time, and mixed space-time comparisons.

The manner in which these comparisons should be made was also discussed at the workshop. It was recognized that a strict one-to-one pairing of predictions and observations at the same location and time is the most severe test to which a model can be put. In view of the stochastic nature of atmospheric turbulence, and hencc diffusion, such a comparison will generally show large differences. The application of models in the regulatory context does not demand close temporal and spatial agreement as the focus is on the maximum concentration values wherever and whenever they occur. The workshop recommended that attention be given to four types of pairing of observations and predictions:

1. The maximum observed value compared against the predicted value for the same time and spatial locations as the maximum observed;
2. The maximum observed and the predicted value at the same time as maximum observed, but anywhere in the spatial domain;
3. The maximum observed and the predicted value at the same location as maximum observed, but at any time;
4. The maximum observed and the maximum predicted value.

Finally, another measure of model performance is to consider the total distribution of observations against a similar display for the predictions. Plotted as a frequency distribution, all the observations and all the predictions can be compared to identify similarity of the distribution and tendencies to predict values that are too large or too small. Similarity of these distributions can be examined by a number of statistical tests. A table that accumulates all of these recommendations is incorporated with Appendix K. We hope the workshop report will be useful in standardizing a set of model performance evaluation techniques.

A few points should be raised with regard to the use of these techniques.

4.3.1.1 Lack of Experience

The list of recommended measures, while being generally accepted as statistical procedures, have only occasionally been applied to air quality model evaluations. A body of experience needs to be developed before we can be sure that the procedure will be valid.

4.3.1.2 Statistical Tests

In a formal statistical context the measures suggested can be tested for acceptance or rejection of the hypothesis that and are equivalent. While this is a useful task, it was not recommended by the workshop because of the inadequacy of observational data and uncertainties associated with the model inputs and parameters.

As Hunter (1978) points out, variance of the observed data

$$s_{c_o} = \frac{1}{N - 1} \Sigma (c_o - \bar{c}_o)^2 ,$$

where N = number of observations, and

$$\bar{c}_o = \frac{1}{N} \Sigma c_o$$

must be sufficiently small with respect to c_o to allow a meaningful comparison with model predictions. If it is not, the difference will be indistinguishable from noise in the data.

Finally, formal testing requires well-behaved data sets, in particular, data must be independent (not serially correlated). As a result of these restrictions the model may be rejected when, in fact, it is a reasonable approximation of the data.

4.3.1.3 Confidence Intervals

In place of formal statistical testing, the workshop group recommended the construction of confidence intervals, using Student's t test for bias and the chi-square test for variance. Confidence statements provide a general indication of the believability of the results.

4.3.2 Scientific Evaluation

The workshop group recognized that a statistical evaluation should be supplemented with a scientific evaluation. One aspect of scientific evaluations involves case-study analyses of interesting data sets in order to discern cause and effect relationships. To perform this type of analysis, more detailed meteorological data are often required than generally needed for model input. Auxiliary data on plume behavior (actual plume rise, measured dispersion rates, etc.) are desirable.

More generally, a scientific evaluation involves examining the merit of the mathematical representations of reality that are in a model. This can be accomplished with a peer review. For the peer review to be meaningful the following characteristics represent minimum requirements.

4.3.2.1 Staff Support for Research

Some group must take the overall responsibility for model review. We do not think that the Office of Air Quality Programs and Studies (OAQPS) should assume this role because of the appearance of, if not actual, conflicts with the regulatory role. While OAQPS must maintain the overall responsibility and authority to issue regulatory guidance, it should not be the reviewer. A major concern of ours with regard to model review has been the final determination in controversial situations. Resolution of such issues requires knowledge, balance, and impartiality, which we believe is most apparently a research function.

Another major concern that we have is the need for follow up of the review. The results of the review should be returned to the model developer so that a process of improvement is initiated. EPA must honor the review and promulgate only models that measure up statistically and scientifically.

4.3.2.2 Five Reviewers as a Minimum

Our experience with peer review suggests that the review quality is directly related to the number of independent reviewers. To ensure that the models receive adequate review, the number of reviewers should be maximized. We consider five reviewers to be a minimum. This number will allow for inadequate reviews or reviews that focus on one particular aspect of the model, and it will reduce the pressure on any one reviewer.

4.3.2.3 Access to Statistical Evaluation

While it is possible to comment on the merits of a model purely on a scientific basis, it is our opinion that much of the scientific basis on which models are built is similar. That which is not common between models is generally more speculative and therefore subject to more controversy. Controversy is best resolved by performance evaluation against data, hence the need for a statistical evaluation becomes greater as the controversy increases. Although one methodology may be clearly more scientifically correct, the translation of that methodology into a model may not lead to improved performance. Hence it is essential that the reviewers have access to the statistical results.

4.3.2.4 Review All Models in a Category

The entire goal of the model evaluation requirement is to select one model over another. This selection process requires that each individual reviewer have access to all the models competing for the same application.

4.3.2.5 Reviewers Must Recognize Controversial Model Issues

Often there has been an accumulation of regulatory action associated with a particular model or modeling procedure (e.g., Class A stability). The reviewers selected to consider a model for applicability should be aware of this controversy. This proposal requires reviewers who are thoroughly familiar with the applications of models and are experienced in regulatory activities. The reviewers should also be unbiased and scientifically qualified. It is a distinct possibility that these two requirements may be mutually exclusive. When in fact it is impossible to obtain reviewers with both qualifications, it is recommended that the number of reviewers be expanded.

4.4 PRINCIPLES OF SELECTION

The Woods Hole Workshop addressed the subject of performance criteria and how good a model must be in order to be recommended for use. The workshop group concluded that it was inappropriate to attempt to set standards at this time. It recognized the desirability of such a simplistic concept but also recognized that standards may be different for different applications and that at present there is no quantitative indication of the accuracy required for various applications.

One approach to establishing criteria of model performance is based on the concept that the model users establish requirements for accuracy and precision. Indeed, this is an important consideration. It is logical that those who use the model and who must make decisions based on its projections have a responsibility to determine what is "good enough" performance. There are, however, a few considerations that complicate this approach to establishing criteria on standards of model performance. First, users are often not in a position to be able to estimate a reasonable degree of model precision and accuracy. Understanding how well a model performs is not a simple task but one that involves a fairly detailed knowledge of meteorology, turbulent diffusion, emissions, and a host of other things. Most users, we expect, would desire a considerably greater degree of accuracy and precision than our current understanding can provide. Second, users of air quality modeling come from a great many different motivations to the model application. Some are engineers concerned about the design of a facility, for example, developing specifications for the required degree of control equipment. Others are regulators concerned about consistency and equity of treatment to various sources. Others are planners requiring perhaps only the most general information. Each requires somewhat different accuracies and precisions. Thus while users must play a major role in the development of model performance criteria, they represent, in our opinion, neither the exclusive nor the primary role. Rather the task is to temper user requirements by the reality of our current and immediately anticipated state of knowledge.

A few considerations that arose concerning discussions of performance criteria bear repeating here.

4.4.1 Scientific Limitations of Models

Atmospheric motions on the scales that act in the transport and diffusion of air pollutants are turbulent. They cannot be predicted or simulated except as some ensemble average over a number of repetitions of the same occurrence. As a result of this random variation there is some inherent uncertainty in any attempt to predict the concentration. This uncertainty decreases as the averaging time increases. In addition, there are uncertainties introduced as a result of variations in the actual emissions. Finally, there are uncertainties associated with the limitations of the data base used in any model evaluation effort. Combined, these lead to a limitation on how well a model can possibly predict concentrations. The model prediction, therefore, should not be considered as a single numerical value but rather as one value within a population of numbers bounded by a probabilistic likelihood. The width of these "confidence limits" or bounds depends on the problem, the meteorological conditions, and a host of other considerations. Decision makers must realize that such bounds exist and that the model prediction is not any tighter than the bounds. Modelers, in turn, must be more diligent in reporting or, at least, in estimating such bounds. While some attempt to assess the confidence limits is recommended, it should also be realized that they may be difficult to estimate.

4.4.2 Comparison with Current Guideline Models

The current models recommended in the Guideline represent (de facto) regulatory acceptability. If a new model can be shown to compare with an acceptable data set better than a guideline model, then the new model should be adopted.

4.5 CURRENT APPROACHES TO MODEL COMPARISONS

The effort to develop a rational basis for model comparisons has generated a number of efforts to conduct the comparisons. Three of these are of sufficient significance to warrant some discussion.

4.5.1 DOE/Savannah River Laboratory

Based on a comprehensive data set relating to emissions from the Savannah River Laboratory, the DOE recently has undertaken an effort to compare various models developed in national laboratories. This effort culminated in a workshop held at Hilton Head, Georgia, in November 1980. Results of this intercomparison are not available at this writing, although the procedures and statistical formats used are outlined in a few preprints from the Atlanta AMS meeting on turbulent diffusion (Pepper, 1981).

4.5.2 Electric Power Research Institute

PRI has undertaken a comprehensive effort to develop a protocol and the necessary data for scientifically documenting the validity of air quality models used in regulatory actions associated with electric power-generating facilities. The program has been documented in a sequence of publications including an initial planning workshop (Hilst, 1978), protocols (Londergan, 1981), statistical evaluation procedures (Londergan, 1980b), and an assessment of existing data bases that could be used for model comparisons (Sklarew, 1979). Many of the statistical measures recommended by AMS (Appendix L) were originally suggested by the EPRI plume model validation effort. Preliminary results of a large field measurement program conducted at a power plant located on relatively smooth level terrain in central Illinois have recently been published (Bowne et al., 1981).

Conclusions reached from the study suggest that the four models tested [(i) Single Source CRSTER (EPA); (ii) MULTIMAX/CEQM (Shell Development Company); (iii) TEM (Texas Air Control Board); and (iv) MSDM (Lawrence Livermore Laboratory) showed significant differences among each other when applied in an operational (standard regulatory) manner.

Preliminary results from the model comparison suggest that 1- and 3-h concentration estimates from models do not compare well with the observations. Cumulative frequency distribution of 1-h maximum concentration as expected compare better than do individual observations. These data are preliminary in nature and will no doubt be broadly discussed and interpreted within the modeling community.

4.5.3 Environmental Protection Agency

The EPA is undertaking efforts to evaluate the performance of models. We commend and support efforts to scientifically quantify and evaluate model performance. We also recognize that decisions suggesting the substitution of one technique over a previously used approach or allowing alternative models must be grounded on the basis of superior performance. A concern exists regarding the appropriate posture with regard to situations where the identification of superior performance is not possible. We are concerned that EPA's press for consistency in such situations may lead to model misuse in particular circumstances and may inhibit the adoption of new, improved models. We believe that scientific judgment should overrule consistency in such instances.

Therefore, we recommend that the EPA refocus its efforts to achieve consistency in modeling applications. The focus should shift away from issues of specific numerical agreement between alternative models and toward issues of consistency of approach to modeling problems. Specifically, EPA seems to interpret the need to specify models "with reasonable particularity" as a mandate that limits the use of scientific merit in the selection of models. We, as representatives of the scientific community, cannot agree to that interpretation. A review of the modeling conference and public hearing records, both of 1977 and 1980, demonstrates considerable sharing of this concern among participants.

EPA's decision to define as recommended models those that passed an equivalency test of 2 percent when compared with the EPA models is an example of the negative aspects of forced consistency. We do not believe that the state of the art of modeling is such that alternative models should be rejected simply because they produce values that differ from those of Guideline models by more than 2 percent. In fact, there is considerable question over the technical validity of the EPA models themselves. Superior models should not be expected to be equivalent. The prescribed procedure for use of an alternative "nonequivalent" model requires that the source gather quantities of on-site data and perform model evaluation that may exceed in effort and quality the information supporting the EPA models themselves.

It is our interpretation that the legislation demands consistency among EPA regional offices in their approach to modeling issues and consistency in the evaluation techniques employed. Consistency in this sense would be in terms of data requirements (amount of data and representativeness), types of dispersion model calculations required (worst-case situation addressed), output requirements (address second highest values, annual averages), and statistical tests employed. One expects, however, to have differences for different regions and applications arising from terrain complexities, effects of local terrain, and other geographic features on dispersion rates, for example. One expects to see legitimate technical agreement and disagreement among models for some situations. It seems that such agreements or disagreements are healthy from the viewpoint of advancing a technology. If resolution of differences cannot be resolved by opposing parties, the EPA Meteorological Laboratory (ORD) and OAQPS can arbitrate. This would be preferred to forcing consistency unilaterally (perhaps imposing a technically unacceptable model on both parties).

The December 14, 1979, opinion of Judge Robinson for the U.S. Court of Appeals for the District of Columbia Circuit addresses the consistency issue and supports the importance of adopting more accurate models in uncertain conditions:

> "Of great importance is a reasoned agency response to substantial questions of fact, policy or science raised in comments on recommended models or in proposals to employ new techniques. In passing Part C of the Clean Air Act, Congress evinced its determination to preserve the clean air regions of the Nation. Congress did not, however, ignore other vital economic and energy considerations. Moreover, successful implementation of the balance struck by Congress will in large part depend on EPA's good sense in establishing and applying modeling guidelines. Modeling, the agency tells us, is on 'the frontiers of scientific knowledge,' but the lack of scientific certitude about modeling techniques increases rather than reduces the need for the agency to critically examine all substantial questions of fact and science emerging from the commenting process. EPA's guideline warns that all proposed deviations from the endorsed model must be fully supported, but this lan-

> guage should not be overemphasized, for the models presently specified in the guideline are concededly flawed. Should scientific advances or better information permit a more accurate assessment of air quality, EPA should move to adopt the more accurate procedure, although it too may not be entirely free from fault."

In conclusion, we believe that the long-term interest of EPA and the modeling community at large will be better served if EPA redirects its efforts toward consistency to methods of problem approach. Through model evaluations, EPA must be relatively free-consistent only with the state of the art-to accept and adopt improvements to the specific models themselves. In short, EPA should attempt to develop a consistent review process that encourages the use of better alternative models.

5

Meteorological and Air Quality Data

In this chapter, we touch on two aspects of the meteorological data problem as related to diffusion modeling--meteorological instrumentation and data collection requirements needed to characterize the airborne dispersion of air contaminants and representativeness of meteorological data. In our opinion, only the first of these topics has been adequately treated as of this date. The importance of air quality data to modeling is also briefly discussed.

5.1 METEOROLOGICAL DATA

The discussions in this section are based on the results of a workshop sponsored by EPA entitled "On-Site Meteorological Instrumentation Requirements to Characterize Diffusion from Point Sources," which was held in Raleigh, North Carolina, January 15-17, 1980 (Hoffnagle *et al.*, 1981, Appendix L). Some 32 scientists drawn from regulatory agencies, industries, universities, research laboratories, and consulting firms participated in the workshop. We endorse the main recommendations stemming from this workshop, which is described in Appendix L.

Meteorological variables to be measured were divided into three priority classes by the workshop participants: (1) essential, (2) desirable, and (3) helpful. These classes reflected a judgment about desirable progress in state-of-the-art modeling and should not limit the modeler in selecting the variables for specific applications. In this report, we will only summarize the "Essential (Priority 1) Measurements." See Appendix L for a more complete discussion.

The workshop results de-emphasize relying on the Pasquill-Gifford-Turner (PGT) stability classification as a primary means of characterizing atmospheric diffusion; rather they emphasize relying on on-site measurements of turbulence intensity. This is, in fact, what Pasquill (1961) recommended when he originally published his scheme. Such a practice, we believe, would represent a major step forward from the techniques used in the 1978 EPA Modeling Guideline. However, considerable model testing and evaluation will be required before the PGT classification scheme can be replaced.

The quantities that have been identified as essential (priority 1) to atmospheric dispersion modeling are as follows:

Mean horizontal wind speed ($\bar{u}$) and mean wind direction ($\bar{\theta}$) at 10 m,
Mean horizontal wind speed and mean wind direction at source height,
Harmonic mean wind speed ($\overline{1/u}$) between source height and effective plume height,
Standard deviations of horizontal and vertical velocity fluctuations (σ_θ, σ_ϕ) at 10 m,
Temperature difference (ΔT) between 2 and 10 m,
Heights of the convectively mixed layer (h_c) and the mechanically mixed layer (h_m),
Average roughness length (z_0),
Friction velocity (u_*),
Bulk Richardson number (Ri_B),
Monin-Obukhov length scale (L),
Turbulent heat flux (H_F),
Convective velocity scale (w_*), and
Water-vapor content (e_w) between source height and effective plume height (if the water content of source emissions is important).

Some of these quantities are primary--direct averages or statistical evaluations of direct measurements--while others are derived from combinations of the primary quantities.

5.2 AIR QUALITY DATA

There is an abundant supply of air quality data available from many sites in the United States. Much of these data are collected by EPA through the National Air Sampling Network and are summarized annually as National Emissions Reports. EPA should review the potential use of this data base as a resource for evaluating air quality models. Using the data, it may be possible to evaluate air quality models under a wide variety of meteorological and geographical conditions. However, generally speaking the data are not perfect or ideally located. Consequently, as with the meteorological data, air quality data contain errors. These errors can lead to inaccurate initial ambient conditions as input for models. Furthermore, source inventories are also estimated with a high degree of imprecision. These errors should be identified and quantified. Decision makers should be made aware that these errors pass through air quality models and contribute to the overall errors in the model simulations.

The accuracy of air quality models is dependent on the precision with which the source can be measured. Moreover, sources of pollution can be divided into three general categories: point, line, and area. The 1978 Guideline document defines these three source categories. In addition, minimum requirements are specified for new source reviews and control strategy evaluations. The Guideline specifies that "for control strategy evaluations the impact of growth on emissions should be considered for the next 10-20 year period."

Site selection for air quality instruments is important to a meaningful determination of maximum ground-level concentrations. Furthermore, evaluations of the impact of planned and existing sources depend upon the sites chosen for air quality instruments. The effectiveness of control strategies are also dependent on proper location of air quality and meteorological instruments. It is important that the location of maximum concentrations (short term and long term) be identified. Instrument fidelity must be maintained.

Background concentrations of regulated effluents must be determined. Representative concentrations for background materials are needed to assess the significance of an effluent release on air quality. The Guideline document addresses procedures for assessing local background concentrations.

5.3 DATA REPRESENTATIVENESS*

Representativeness is a term meteorologists use to describe a desirable attribute of the data they apply and the results of analyses they perform, but it is not easy to find a definition for the term in meteorological literature. This is important for deciding the location and duration of data for air quality analyses. The term, for example, does not appear in the Glossary of Meteorology, published by the AMS (Huschke, 1950), although it does appear in the Meteorological Glossary, published by the British Meteorological Office. Petterssen (1940) in his well-known early text on weather analysis and forecasting enjoins the fledging meteorologist "to ascertain that the observations which we make use of are representative. By this we mean that the observed quantity should be representative of the value characteristic of the entire air mass or a large portion thereof. This is the same as saying that the observations should not be influenced locally." Others, writing on weather forecasting, have asserted that a representative observation is one that can be applied to synoptic weather analysis. If it cannot be so applied it is "irrepresentative" (Godske et al., 1957). Such a definition for "representative" appears less appropriate for meteorological matters and air quality determinations pertaining to smaller scales. Instead of integrating a number of simple observations to produce a datum representative of a large area there is a need to specify the area to which a single observation may be usefully applied (Geiger, 1950; Zemel and Lyons, 1976).

Geiger in his classic editions of The Climate Near the Ground (1950; 1961) defines a representative station as "a station which corresponds to the average relationships between its nearby and more distant surroundings." He argues that "the chief requirement of a representative station is that it shall have a wide range of validity." He speaks of the "range of validity" as a feature of the numerical data of an observation station being considered applicable to a wider territory than its immediate vicinity. The range of validity would end where the numerical departures from the station's data become too great to be neglected on the basis of the accuracy required. More specifically, if two stations produce data whose differences are not random, they are not in the same range of validity. The two stations could be said to be unrepresentative or "irrepresentative" of each other.

---

*The original material for this section was prepared by Herschel H. Slater, formerly with EPA, Research Triangle Park, NC.

Ball and Anderson (1977) in developing criteria for siting air pollution monitors define a "representative spatial scale" by following Geiger's arguments. A representative spatial scale is "the distance from a monitoring station at which measurements become significantly different from those at a monitoring station."

Data that for some purposes have the character of undesirable noise are for other purposes highly relevant. An observation of temperature collected in a deep valley may be unrepresentative for synoptic purposes but essential and highly relevant for predicting local frost (Godske, 1966). Representativeness of data is influenced by the purpose for which the data are used. Similarly, representativeness may differ for different weather elements.

For air quality management activities there are stringent requirements for representative weather data. A major problem in obtaining representative data is that traditionally weather data have been sampled regularly in time but erratically in space (Baer and Tribia, 1976). Spatial irregularity has always been burdensome to the synoptic meteorologist and probably even more so to the air pollution meteorologist. The detailed weather information necessary to evaluate properly the environmental impact of a proposed source of pollution is almost never available at the time critical decisions as to the source's location, size, and operating characteristics are made. A frantic search for representative weather data ensues.

Measures of representativeness are those that allow the analyst to judge the degree of similarity of the data sets collected at two or more sites. A common statistic for this purpose is the spatial autocorrelation (Zemel and Lyons, 1976). Another is some test of the similarity of cumulative frequency distributions of weather elements (Turner, 1974). In both cases there is no clear procedure for determining the criteria for judging whether the similarity is sufficient to conclude that two or more data sets are representative of each other. However, procedures can be used to identify the locations that have relatively greater similarity (and are presumably more representative of each other) than others. Prunchinicki (1975) uses probability distribution functions to consider stations alternately and arrives at an unbiased selection of a representative station. Details are sketchy, but he developed and applied an index of similarity. Whether the degree of similarity is sufficient to meet the needs of a particular problem is unclear.

The need for representative meteorological data is uniformly recognized. The measures to characterize the representativeness of a data set are not widely agreed on. Neither are procedures to establish values that characterize adequate representativeness. As Hayes (1979) discusses with respect to choosing performance standards, it appears appropriate to resort to the pragmatic-historic approach to arrive at the needed decision. Objective measures are applied and the results documented wherever possible. Redundancy of data should not be avoided.

The analyst must now assess the objective and comparative data available and make a judgment. He must answer, "Are the data available sufficiently similar to be considered usable for the problem at hand?" He then has harder questions, "Are the data at hand sufficiently similar to what would be observed at a designated location for which data are lacking?" It is likely that the objective data are not adequate to allow the analyst to state with confidence that the data are usable. The analyst must apply other subjective tests. He must take great care that they are founded on physical premises. He must avoid the appearance or implication of applying professional indulgence to dismiss available data. However, he is responsible for investigating the manner in which the data were collected, the quality assurance procedures that were exercised, and the competence of those who performed these tasks. He must assess the exposure of the sensors and of the setting. Were the data available for analysis collected under essentially the same conditions as exist at the site in question (Hayes, 1979)? Are the gradients of values of meteorological elements weak between the two sites? Are the macroclimate features similar? Are the radiation relationships alike (Geiger, 1950)? Does the terrain slope the same? Are horizons similar? If not, are differences sufficient to affect adversely the application of the data? Is the degree of similarity dependent on macroscale features (Munn, 1975)?

If these and other questions are consistently studied and responded to and if weaknesses and strengths of the analysis are documented and openly discussed, then agreements on the representativeness of data for specific air quality management problems can be reached.

References

Baer, F., and J. J. Tribia (1976): Spectral fidelity of Gappy data, *Tellus 27*, 215-226.

Ball, R. J., and G. E. Anderson (1977): Optimum site exposure for SO_2 monitoring. EPA-450/3-77-013, Office of Air Quality Planning and Standards, Environmental Protection Agency, Research Triangle Park, NC.

Bass, A. (1980): Modeling long range transport and diffusion. *The Second Joint Conference on Applications of Air Pollution Meteorology* (Conference Papers), American Meteorological Society, Boston, MA, pp. 193-215 (see Appendix E of this report).

Bowne, N. (1980): Validation and performance criteria, air quality models. *The Second Joint Conference on Applications of Air Pollution Meteorology* (Conference Papers), American Meteorological Society, Boston, MA, pp. 614-626 (see Appendix F of this report).

Bowne, N. (1981): *Preliminary Results from the EPRI Plume Model Validation Project--Plains Site*, Electric Power Research Institute, Report No. EA-1788-SY, TRC, 125 Siles Dean Hwy, Wethersfield, CT 06109.

Box, G. E. P., and G.M. Jenkins (1976): *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, CA, 575 pp.

Briggs, G. A. (1975): Plume rise predictions, in *Lectures on Air Pollution and Environmental Impact Analyses*, Workshop Proceedings, American Meteorological Society, Boston, MA, pp. 59-111.

Briggs, G. A., and K. R. McDonald (1978): Prairie Grass revisited: optimum indicators of vertical spread, in *Proceedings of the NATO/CCMS 9th International Meeting on Air Pollution Modeling and Its Application*, available from U.S. Atmospheric Turbulence and Diffusion Lab., Oak Ridge, TN, Contribution 78/8, June 1978, 31 pp.

Calder, K. L. (1949): Eddy diffusion and evaporation in flow over aerodynamically smooth and rough surfaces, Quart. J. Mech. Appl. Math. II 153.

Clarke, R. H. (1979): A model for short and medium range dispersion of radionuclides released to the atmosphere. National Radiological Protection Board, NRPB-R91, Harwell, Oxon, UK, 74 pp.

Dickerson, M. H., and P. H. Gudicksen (1980): The DOE's Atmospheric Studies in Complex Terrain (ASCOT) program, The Second Joint Conference on Applications of Air Pollution Meteorology (Conference Papers), American Meteorological Society, Boston, MA, pp. 467-473.

Donaldson, C. du P. (1973): Atmospheric turbulence and the dispersal of atmospheric pollutants, Workshop on Micrometeorology, D. A. Haugen, ed., American Meteorological Society, Boston, MA, pp. 313-390.

Egan, B. A. (1975): Turbulent diffusion in complex terrain. Lectures on Air Pollution and Environmental Impact Analyses (D. Haugen, ed.), American Meteorological Society, Boston, MA, pp. 112-135.

Egan, B. A., and J. R. Mahoney (1972): Numerical modeling of advection and diffusion of urban area source pollutants, J. Appl. Meteorol. 11, 312-322.

Environmental Protection Agency (1977a): Conference on Air Quality Modeling, Acme Reporting Company, Washington, D.C. 20005, December 1977.

Environmental Protection Agency (1977b): Guideline Vol. 10, Reviewing New Stationary Sources, EPA-450-14-77-00.

Environmental Protection Agency (1978): Guideline on Air Quality Models, EPA-450/2-78-027.

Falls, A. H., and J. H. Seinfeld (1978): Continued development of a kinetic mechanism for photochemical smog, Environ. Sci. Technol. 12, 1398-1405.

Fox, D. G. (1981): Judging air quality model performance--review of the Woods Hole workshop, in Proceeding of Fifth Symposium on Turbulence, Diffusion, and Air Pollution, American Meteorological Society, Boston, MA.

Fox, D. G., and J. E. Fairobent (1981): NCAQ Panel examines uses and limitations of air quality models, Bull. Am. Meteorol. Soc. 62, 218-221.

Fox, D. G., and S. A. Orszag (1973): Pseudospectral approximation to two-dimensional turbulence, J. Comput. Phys. 11, 612-619.

Geiger, R. (1950): The Climate Near the Ground, 2nd ed., Harvard U. Press, Cambridge, MA.

Geiger, R. (1961): The Climate Near the Ground, 4th ed., Harvard U. Press, Cambridge, MA.

Gifford, F. A. (1981): Horizontal diffusion in the atmosphere: a Lagrangian-dynamical theory, Atmos. Environ. 15 (in press).

Godske, C. L. (1966): Methods of statistics and some applications to climatology, in *Statistical Analysis and Prognosis* in *Meteorology Tech Note*, No. 71, WMO-No. 178, TP.88, WMO, Geneva, Switzerland.

Godske, C. L., T. Bergeron, J. Bjerknes, and R. C. Bundgaard (1957): *Dynamic Meteorology and Weather Forecasting*, American Meteorological Society, Boston, MA.

Hanna, S. R. (1981): Lagrangian and Eulerian time-scale relations in the daytime boundary layer, *J. Appl. Meteorol. 20*, 242-249.

Hanna, S. R., G. A. Briggs, J. Deardorff, B. A. Egan, F. A. Gifford, and F. Pasquill (1977): Summary of recommendations made by the AMS Workshop on stability classifications schemes and sigma curves, *Bull. Am. Meteorol. Soc. 58*, 1305-1309.

Hayes, S. R. (1979): Performance Measures and Standards for Air Quality Simulation Models, EPA-450/4-79-032, Office of Air Quality Planning and Standards, Environmental Protection Agency, Research Triangle Park, NC, 160 pp.

Heffter, J. L., A. D. Taylor, and G. J. Ferber (1975): A Regional-Continental Scale Transport, Diffusion, and Deposition Model, *NOAA Tech. Memo* ERL-ARL-50, 28 pp.

Hicks, B. B., M. L. Wesley, and J. L. Durham (1980): Critique of Methods to Measure Dry Deposition--Workshop Summary, EPA-600/9-80-050, EPA, Washington, D.C.

Hillyer, M. J., S. D. Reynolds, and P. M. Roth (1979): Procedures for evaluating the performance of air quality simulation models, EPA-450/4-79-033, Environmental Protection Agency, Research Triangle Park, NC, 146 pp.

Hilst, G. R. (1978): *Plume Model Validation*, Workshop Summary Report EA-917-SY, Electric Power Research Institute, Palo Alto, CA.

Hoffnagle, G. F., M. E. Smith, T. V. Crawford, and T. J. Lockhart (1981): On-site meteorological instrumentation requirements to characterize diffusion from point sources--a workshop. *Bull. Am. Meteorol. Soc. 62*, 255-261.

Hosker, R. P. (1980a): Practical application of air pollutan deposition models, in *Proceedings of the International Conference on Air Pollutants and Their Effects on the Terrestrial Ecosystem*, S. V. Krupa and A. H. Legge, eds., Wiley, New York.

Hosker, R. P. (1980b): Dispersion in the vicinity of buildings. *The Second Joint Conference on Applications of Air Pollution Meteorology* (Conference Papers), American Meteorological Society, Boston, MA, pp. 92-107 (see Appendix G of this report).

Hunter, J. S. (1978): Appendix A, Plume model validation, EA-917-SY, Workshop WS-78-99 Summary Report, October 1978, Electric Power Research Institute, Palo Alto, CA.

Huschke, R. E., ed. (1950): *Glossary of Meteorology*, American Meteorological Society, Boston, MA.

Interagency Task Force on Acid Precipitation (1981): *The National Acid Precipitation Assessment Plan*, CEQ, January draft 129 pp., available from Council on Environmental Quality, 722 Jackson Place, NW, Washington, D.C. 20006.

Irwin, J. S. (1979): Estimating plume dispersion--a recommended generalized scheme, in *Proceedings of the Fourth Symposium on Turbulence, Diffusion, and Air Pollution*, American Meteorological Society, Boston, MA, pp. 62-69.

Johnson, W. B. (1980): Interregional exchanges of air pollution: model types and applications, *Proceedings of the 11th Technical Meeting on Air Pollution Modeling and Its Application*, NATO, Amsterdam, The Netherlands, 24-27 Nov. 1980.

Johnson, W. E., R. C. Sklarew, and D. B. Turner (1976): Urban air quality simulation modeling, in *Air Pollution*, A. C. Stern, ed., Vol. 1, Chap. 10, pp. 503-562.

Kaimal, J. S., J. C. Wyngaard, D A Haugen, O. R. Cote, Y. Izumi, J. J. Caughey, and C. J. Readings (1977): Turbulence structure in the convective boundary layer, *J. Atmos. Sci. 33*, 2152-2169.

Lamb, R. G. (1979): The effects of release height on material dispersion in the convective planetary boundary layer. *In Proceedings of the Fourth Symposium on Turbulence, Diffusion, and Air Pollution*, American Meteorological Society, Boston, MA, pp. 27-33.

Latimer, D. A., R. W. Bergstrom, C. D. Johnson, and J. P. Killus (1980): Modeling visibility, *Second Joint Conference on Applications of Air Pollution Meteorology*, American Meteorological Society, Boston, MA, pp. 346-360 (see Appendix H of this report).

Lewellen, W. S., and Y. P. Sheng (1980): *Modeling Dry Deposition of SO_2 and Sulfate Aerosols*, EA-1452 Research Proj. 1306-1, Electric Power Research Institute, Palo Alto, CA.

Londergan, R. J. (1980): *Protocol for Plume Model Validation*, EPRI EA-1638, Electric Power Research Institute, Palo Alto, CA.

Londergan, R. J. (1981): *Validation of Plume Model Statistical Methods and Criteria*, EPRI EA-1673-SY, Electric Power Research Institute, Palo Alto, CA.

Lyons, W. A. (1975): Turbulent diffusion and pollutant transport in shoreline environments, Lectures on Air Pollution and Environmental Impact Analyses, D. Haugen, ed., American Meteorological Society, Boston, MA, pp. 136-208.

McCutchen, M. H., D. G. Fox, and R. W. Furman (1981): The conical mountain study--determining the effect of elevation and aspect, preprint volume, Second Conference on Mountain Meteorology of the American Meteorological Society, Boston, MA, pp. 75-81.

McElroy, J. L., and F. Pooler (1968): St. Louis Dispersion Study, U.S. Public Health Service, National Air Pollution Control Administration, Report AP-53.

McMahon, T. A., and P. J. Denison (1979): Empirical atmospheric deposition parameters--a survey. Atmos. Environ. 13, 571-585.

Munn, R. E. (1975): Suspended particulate concentrations: spatial correlations in the Detroit-Windsor area. Tellus 27, 397-405.

Orgill, M. M. (1981): Atmospheric Studies in Complex Terrain: A Planning Guide for Future Studies, PNL-3656, UC-11, ASCOT 80-4, Pacific NW Labs, Richland, WA.

Pasquill, F. (1961): The estimation of the dispersion of windborne material, Meteorol. Mag. 90, 33.

Pasquill, F. (1974): Atmospheric Diffusion, Halsted Press, New York.

Pasquill, F. (1976): Atmospheric Dispersion Parameters in Gaussian Plume Modeling: Part II. Possible Requirements for Change in the Turner Workbook Values, EPA-600/4-760306, Environmental Protection Agency, Research Triangle Park, NC, 44 pp.

Pepper, D. W. (1981): Results from the Savannah River Laboratory model validation workshop, Fifth Symposium on Turbulence, Diffusion, and Air Pollution (Extended Abstracts), American Meteorological Society, Boston, MA, pp. 12-13.

Petterssen, S. (1940): Weather Analysis and Forecasting, McGraw-Hill Book Co., Inc., New York.

Pielke, R. A. (1974): A three-dimensional numerical model of the sea breezes over south Florida, Mon. Wea. Rev. 102, 115-139.

Prahm, L. P., and O. Christensen (1977): Long-range transport of pollutants simulated by a two-dimensional pseudospectral dispersion model, J. Appl. Meteorol. 16, 896-910.

Prunchinicki, J. (1975): Similarity index as a measure of representativeness of meteorological stations (English translation of title), Przeglad Geofizyczny Warsaw 20, 319-324.

Randerson, D. (1981): *Atmospheric Science and Power Production*, U.S. Department of Energy, Technical Transfer Division, Oak Ridge, TN (in press).

Reid, J. D. (1979): Markov chain simulations of vertical dispersion in the neutral surface layer for surface and elevated releases, *Boundary Layer Meteorol. 16*, 3-22.

Roberts, J. J., ed. (1977): *Report to the U.S. EPA of the Specialists' Conference on the EPA Modeling Guidelines*, Environmental Protection Agency, Research Triangle Park, NC.

Ruff, R. E. (1980): Evaluation of the real-time air-quality model using the RAPS data base, EPA-600/4-80-013a, Vol. I, Overview, by SRI International, Menlo Park, CA, 22 pp.

Sehmel, G. A. (1981): Deposition and reinsertion processes, Chap. 12 in *Atmospheric Sciences and Power Production*, D. Randerson, ed. (to be published by U.S. Department of Energy).

Sehmel, G. A. (1980): Particle and gas dry deposition: a review. *Atmos. Environ. 14*, 983-1011.

Sherman, C. A. (1978): A mass consistent model for wind fields over complex terrain, *J. Appl. Meteorol. 17*, 312-319.

Sklarew, R. (1979): *Data Bases for Plume Model Validation*, EPRI EA-1159, Electric Power Research Institute, Palo Alto, CA.

Slade, D. H. (1968): *Meteorology and Atomic Energy*, USAEC Division of Technical Information Extension, Oak Ridge, TN.

Slater, H. H. (1977): *Comments and Recommendations Concerning the Draft Guidelines on Air Quality Models*, Environmental Protection Agency, Research Triangle Park, NC.

Slinn, W. G. N., L. Hasse, B. B. Hicks, A. W. Hogan, D. Lal, P. S. Liss, K. O. Munnich, G. A. Sehmel, and O. Vittori (1978): Some aspects of the transfer of atmospheric trace constitutents past the air-sea interface, *Atmos. Environ. 12*, 2055-2087.

Thom, A. S. (1975): Momentum, mass and heat exchange of plant communities, *Vegetation and the Atmosphere*, J. G. Monteith, ed., VZ Academic Press, New York, pp. 57-109.

Trout, D. A. (1979): A review of horizontal and vertical dispersion coefficients applied in the calculation of ground level pollutant concentrations resulting from emissions from tall stacks under very unstable atmospheric conditions, in *Proceedings of the NATO/CCMS 10th International Technical Meeting on Air Pollution Modeling and Its Applications*, Oct. 23-26, 1977, Rome, 10 pp.

Turner, D. B. (1974): Frequency distributions from dispersion models, in Proceedings of the Symposium of Statistical Aspects of Air Quality Data, pp. 10-1-10-20, EPA-650/4-74-038, Environmental Protection Agency, Research Triangle Park, NC.

Venkatram, A., 1980a: The relationship between the convective boundary layer and dispersion from tall stacks, Atmos. Environ. 14, 763-767.

Venkatram, A., 1980b: Dispersion from an elevated source in a convective boundary layer, Atmos. Environ. 14, 1-10.

Weil, J. (1979): Assessment of plume rise and dispersion models using lidar data, Martin Marietta Corp., Environ. Center, 1450 S. Rolling Blvd., Baltimore, MD, prepared for the Maryland Power Plant Siting Program, Report No. PPSP-MP-24.

Wilson, N. R., and R. H. Shaw (1977): A higher-order closure model for canopy flow, J. Appl. Meteorol. 16, 1197-1205.

Zemel, Z., and J. Lyons (1976): An objective method for assessing representativeness of a station network measuring minimum temperature near the ground, Boundary-Layer Meteorol. 10, 3-14.

Appendix A

SUMMARY REPORT ON EPA'S 1977 AIR QUALITY MODELING CONFERENCES

by

Gordon A. Beals, Ph.D
Member, AMS Committee on the
Meteorological Aspects of Air Pollution

Prepared for the Steering Committee
AMS-EPA Cooperative Agreement

December, 1980

SUMMARY REPORT ON E.P.A.'S 1977 AIR QUALITY MODELING CONFERENCES

The Environmental Protection.Agency (EPA) held a series of five public meetings between May 17, 1977 and June 1, 1977 for the purpose of presenting its draft Guidelines on Air Quality Models and receiving comments from interested members of the public. The stated aim of the guidelines was to promote consistency, if not uniformity, in the application of air quality models. The transcipts of the meetings contain comments from approximately 135 individuals who were not EPA employees. These represented approximately 125 state and local pollution control agencies, industries, consultants, and other interests. There were also about 50 sets of writters comments.

Comments were received on a wide range of topics, including five on which EPA specifically solicited advice. The five questions specifically asked by EPA were:

1. What operating level for point sources should be used when ascertaining probable impact on air quality standards?

2. What length of meteorological record is sufficient for air quality model estimates and what short-term estimate should be the basis for emissions regulations?

3. How are background air quality data best determined?

4. To what extend should specific analytical techniques be required for assessing the impact of a specific source or source categories?

5. How frequently should the guideline be revised?

The last topic received approximately 30 comments. Most of the commentors preferred annual revision. Some made a distinction between review and revision, saying the guideline should be reviewed at least annually, but revised only as necessary to include significant new developments. Six participants suggested longer revision intervals from 14 months to 5 years. The motivation for less frequent review was to create a period of stability long enough to complete the source rewiew process without changing models in mid-stream. This concern for continuity was also expressed by several persons who advocated frequent revision. They suggested that the same model should be used throughout the review process for a given project and the modeling should be "grandfarthered" against retroactive evaluations based on new models introduced by a guideline revision.

Several commentors used the revision theme to comment about the purposes and organization of the guideline, and on the method of revision. Six or more commentors suggested that the guideline be organized into a policy section and a "handbook" section, the latter section perhaps being in looseleaf form. The handbook part could then be revised rapidly to reflect the latest thinking on a particular technique while the policy part would be revised much less frequently. As to the method of revision, some commentors felt that the specialist conference format was appropriate. Two commentors, one a state agency, the other an industrial spokesman, requested use of a formal revision-announcement public-meeting procedure for every revision.

Forty comments were received directly relating to the assumed operating level for ascertaining probable impact on air quality standards. Persons whose primary concern was process emissions generally suggested using average emission rates for modeling process sources but suggested rated load for all boiler evaluations. One commentor requested that the guideline discuss the need (or absence of need) to model cases of abnormal or "upset" emissions. For power plants, one commentor suggested that an assumption of 100% load for all averaging times would be appropriate for "baseload plants". Several commentors suggested using annual average load for evaluating compliance with long term air quality standards, but using maximum load for evaluation of compliance with short term standards, in the case of single-source modeling. Most of those advocating the use of maximum load for short term modeling also expressed a preference for assuming "reasonable worst case" meteorology. Assumption of simultaneous occurrence of maximum load together with extreme worst case meteorology should be avoided. Several suggested doing the short term evaluation two ways; first worst case load with average meteorology and average load with worst case meteorology. Two state agencies stated that they usually required evaluation at two or more load levels; maximum, most probable, and minimum. One utility commentor suggested that three levels was too many cases because maximum load is usually the worst case. Others felt that the maximum load case is not necessarily the worst. One pointed out that, because plume buoyancy is correlated with load, there might be little difference between the impacts at different load levels.

Eight commentors felt that diurnal load variations were important enough to be considered in 24-hour modeling. If multiple years of meteorology were used, it was suggested by one that the same emission pattern should be used for all years. Seven commentors suggested statistical approaches to ascertaining short term impact, requiring that the joint probability of meteorology and load be considered. The approach suggested here was similar to EPA's "Ex-Ex" method of handling coal sulfur variability. Those who did not seek such a sophisticated model felt that at least the joint probability of occurrence of maximum load and maximizing meteorology should be evaluated.

The simultaneity problem is compounded in the case of short term multi-source models. It was pointed out that there would be an extremely low probability of all sources operating simultaneously at maximum load.

More than 70 comments were received concerning the length of meteorological data record and the representativeness thereof. Nearly half addressed length of record only, while the others addressed various aspects of representativeness. Most of the commentors in the first group agreed that five years is a desirable length of record to obtain a good assessment of the probability of occurrence of adverse meteorological conditions. However, many of these same commentors felt that a 5-year period was impractical and too costly. It was claimed that hourly calculations over a five-year period requires computational and manpower resources beyond those available to many model users. Furthermore, the desired meteorological data were often not readily available from the National Climatic Center. Several commentors in this group made a distinction between NCC data and on-site data, claiming only that five years of on-site data gathering would be too burdensome. They recommended the use of one year if on-site data were used.

Some commentors felt that the decision on length of record should be made on a case by case basis in conference between a permit applicant and the regulatory agencies involved. They noted that the complexity and expense of data acquisition depended on the type of model being applied. They preferred that the guideline focus on criteria for assessing the adequacy of the data base.

One requested a detailed listing of the input data requirements for each model. A few commentors asked for guidance on estimating the frequency and duration of adverse conditions as an alternative to sequential hourly modeling.

Six commentors specifically questioned the utility of NCC data for any modeling. Two of these (one an East Coast utility, the other a West Coast regulatory agency) mentioned the "instantaneous" nature of National Weather Service observations as distinct from true hourly averages. One person suggested that recent NCC data, based on observations taken only every third hour, should not be used for modeling maximum concentrations for a three-hour averaging period. The others questioned the representativeness of airport data even a few miles from the site of source being modeled. Three commentors, while not specifically doubting the utility of off-site data, stated a need for criteria to judge whether off-site data are adequate.

Several commentors suggested that it would be possible to manufacture a meteorological record for a specific site by combining a short period of meteorological data from the site itself with a longer period of data from some nearby airport. These commentors felt that at least the guideline should provide some suggestions on how to accomplish such a merger. Others pointed out that meteorological data are really necessary in three dimensions and that three-dimensional data are extremely difficult to obtain. In response to this need, another commentor suggested that there should be a systematic program of gathering data in the lowest thousand meters of the atmosphere. This could be a function either of the National Weather Service or EPA. Finally, some participants claimed that the uncertainties inherent in the models and other aspects of the data are such that a five year period of record provides no more useful information than a one year period.

In response to EPA's questions, there were about forty comments on best methods to determine background concentrations. Several commentors requested that the meaning of the term "background" be clarified. The participants showed a strong preference that background or baseline air pollution concentration data be determined by analysis of air quality measurements. However, many of these same participants also recognized that modeling would be valuable as a means of extrapolating or interpolating between concentrations measured at monitoring stations. Some of the commentors expressed the opinion that modeling done for this purpose is best left to the air pollution control agencies. Some of these participants also cautioned that if modeling revealed any so-called "hot spots" that these should not be regarded as necessarily the correct maximum air quality concentrations until these "hot spots" had been investigated by monitoring.

One commentor mentioned special problems associated with determining background in the presence of existing sources. He pointed out that there is a danger that, unless one has a very well calibrated model for a particular source, one might conclude that too much background would be removed if emissions from that particular source were reduced or eliminated. He also wondered how emission reductions in general should be treated in determination of PSD background.

A number of commentors offered specific suggestions for determining background. One suggested that background was stability-dependent and another suggested that there was a dependence on wind speed such that one ought to use a curve of background versus wind speed to estimate the background concentration applicable to a given modeled increment prediction. Several others suggested the use of pollution wind roses. Six of the commentors suggested that if an actual sequence of meteorology is used in the modeling, then the corresponding sequences of measured air quality data should be used as background values.

The consideration of future growth also received four comments under this topic. The concensus seemed to be that modelers should consider all the emissions out to the most future date for which they have any information. However, one commentor stated that he doubted anyone's ability to forecast growth twenty years ahead.

Almost all of the other comments can be thought of as responding to the question "To what extent should specific analytic techniques be required for assessing the impact of specific sources or source categories?" There were numerous comments on the general issue of flexibility versus specificity. A great majority of the commentors preferred a flexible approach. One or two even felt that the guideline in its present form had little utility. Several of the commentors who preferred a flexible approach expressed a fear that the guideline as it now stands would be interpreted inflexibly, more like a regulation. Some of these participants felt that almost any departure from the guideline should be permitted, provided that supporting documentation on the alternative model's performance was supplied to the Regional Administrator. One participant specifically suggested that the guideline should say that models and techniques not specifically mentioned in the guideline should not be regarded as disapproved. Five or six of the comments suggested that the guideline should address situations to be avoided rather than recommending specific models. They thought the guideline should perhaps consist of a catalog of different techniques with some strictures against applying certain models in certain situations, and that the guideline should contain a checklist of factors to be considered when applying different models. Some commentors suggested that the guideline should lay out criteria for accepting alternative models. Another requested that the guideline define what would be a minimum acceptable modeling effort in various situations. There was strong support for the concept of case-by-case modeling decisions wherein air pollution control officials at various levels would meet to discuss modeling techniques, or perhaps in which a perspective source constructor would meet with the regulatory officials to decide on the specifics of the modeling approach. Eight of the commentors apparently felt that EPA would not be flexible enough. These commentors suggested that some sort of review mechanism be established external to EPA.

Some of those who expressed at least some preference for standardization preferred that algorithms be standardized rather than computational codes, as the EPA models are. The most common reason for this preference was to optimize computation.

A few of the commentors addressed the issue of consistency, which is believed to be the primary motivation for EPA issuing its modeling guideline. Some said that there definitely should be consistency from state to state and from EPA region to EPA region. This is desirable so that the review process would be essentially the same from region to region. Others saw less need for consistency between EPA regions but felt it highly desirable that there should at least be consistency within the same region. One participant stated that the consistency that the iron and steel industry was looking for is a consistency from region to region in the procedure for making modeling decisions, not necessarily using the same models in different locations. A few of the participants pointed out that some EPA modeling recommendations appeared to be internally inconsistent. For example, the Climological Dispersion Model with statistical adjustment and RAM are both recommended for short term urban modeling, but give very different results for the same emissions. Another perhaps, less extreme inconsistency, was the recommendation either RAM or CRSTER. This was felt to be inconsistent because these otherwise similar models have different sigma values. Although most participants prefer to avoid standardization of models, one suggested the concept of a reference model wherein EPA would establish what meteorological phenomena are important to treat explicitly and then propose a reference algorithm to treat each of these elements.

There was a rather striking exception to this general feeling that the guideline was too specific and needed more flexibility. The reversal concerned the more complex issues. All in all, some 25 participants suggested that the guideline should provide more guidance on several of the difficult issues, including complex terrain, stack and aerodynamic downwash, and fugitive particulates including the problem of settling particulates. Several of the people who would have liked more guidance on complex terrain recognized the knowledge gap and requested that EPA devote high priority in its research program to developing dispersion models applicable to complex terrain. Two commentors felt that perhaps the guideline would be useful if it indicated criteria indicating when the Gaussian models clearly should not be used for complex terrain. There were a few affirmative comments about the complex terrain modeling. Specifically,one participant suggested that the wind field representation and the dispersion coefficient representation be separate submodels when modeling complex terrain. Another participant suggested using the simple box model with an inversion lid.

Another tough issue which came up frequently in conjunction with complex terrain was long range transport. Two general types of comments were made. First was a request that removal processes be considered in long range transport models, and the second was that some maximum downwind distance be set for Gaussian models. The suggested limitations for downwind distance ranged from 25 kilometers to a statement "much less than 100 kilometers." One participant noted that there was nothing factual about sigma y and sigma z beyond distances of about 10 to 15 kilometers. Several commentors felt that it would be appropriate for the guideline to give somewhat more emphasis to other types of modeling such as physical modeling and numerical models. In connection with trajectory modeling, it was recognized that the application of these models require a much more extensive data base than the application of Gaussian models. It was felt that either this guideline or some other publication should point out and discuss such issues as whose responsibility it is to plan, organize, and pay for the data gathering effort.

In the San Francisco meeting there were approximately 20 comments addressing the subject of photochemical oxidant modeling. Many of the commentors were dissatisfied with what EPA calls the isopleth model. Some of the commentors suggested that photochemical trajectory modeling should be included in the guideline along with such items of guidance as criteria for selecting a suitable model and specification of necessary data bases.

Another topic where a specific technique was advocated by a number of participants was the use of statistical techniques for estimating short-term concentrations. Two participants cautioned that the Larsen technique for extrapolation of concentrations to different averaging times is not valid for large point sources. Another stated the converse of this, i.e. that the technique should be applied only to urban multi-source situations. For the most part this large group of commentors supported the concept of looking at the statistical distribtuion of concentrations at each modeled receptor. Most of this group proposed what the statisticians would call a parametric approach rather than an enumerative approach to find the concentration that would be exceeded more than once per year. This kind of procedure would obviate the suggestion in the guideline that the highest concentration estimate be used when the data base was considered to be inadequate. The second highest concentration at the point would be estimated based on the portion of the frequency distribution where the data was adequate.

A related question raised by one of the state agencies is what if the short term high concentration estimate were based on a synthesized worst case meteorological estimate instead of a sequence of actual meteorological measurements? What frequency does one assign to this synthesized worst case condition? Also related to the short term concentrations issue was a suggestion that ten-minute to thirty-minute wind averages be used instead of sixty-minute wind averages when available. That is that two to six periods or subperiods be used to represent a single hour.

There were also a large number of comments concerning the specific EPA recommended modeling techniques rollback, VALLEY and CRSTER. The rollback technique was almost universally regarded as not an appropriate model. However one commentor did point out certain carefully constrained situations where rollback modeling would be appropriate. With regard to VALLEY it was felt the model should be used for screening or not at all. The most common criticism of CRSTER concerns the treatment of calm periods by the meteorological preprocessor. The commentors felt that assuming a persistent wind direction at one meter per second for successive hours of calm was not a reasonable procedure. However, no one offered a substitute. Two other commentors felt that representing an entire hour by a single airport wind direction measurement was not appropriate because, in light wind situations, the wind direction fluxuations within an hour may be larger than 10 degrees.

In the case of stable plumes this is also much larger than the width of the plume. Therefore, some sort of additional randomization of direction within the hour seemed appropriate.

The next major topic with respect to CRSTER is the terrain clearance issue. Several commentors suggested that the procedure adopted in VALLEY, that is nearly full terrain subtraction in the case of stable hours but no terrain subtraction in the case of neutral or unstable hours, was more appropriate than full terrain subtraction for all hours, as is done in CRSTER. Some others expressed a preference for the Egan half-terrain correction. One commentor who had used aircraft to track plumes in complex terrain stated that he was convinced that full terrain subtraction was a better representation than others in stable conditions. Finally there were some misgivings expressed about the method of interpolating inversion heights used in the CRSTER or RAMMET meteorological preprocessors.

Two persons noted apparent underpredictions by the CRSTER model in certain meteorological conditions. Another suggested that the guideline should be more explicit as to what meteorological and terrain complexities would rule out the use of CRSTER.

The RAM urban model was discussed by two or three commentors. The reported experience with RAM was almost universal overprediction as far as had been measured.

The topic of plume rise formulas was frequently mentioned in the context of the CRSTER model. Several commentors noted that there were more than one equation that could be described as the Briggs plume rise equation, and requested that EPA clarify which Briggs reference they mean when prescribing the Briggs plume rise formula. Several other participants suggested that plume rise enhancements caused by combined buoyancy from stacks in close proximity to each other should be addressed in the guideline. TVA suggested that, for tall stacks, somewhat smaller gradients of potential temperature with height be assumed for stable plume rise calculations. One participant requested that EPA do more work on verifying plume rise from small sources. Some participants also requested guidance on handling of elevated inversions above stack top, both with respect to the reduced plume rise into the elevated inversion, and with respect to the issue of partial penetration.

Many participants took the opportunity to recommend validation or calibration of models. In general most of the participants who commented on this issue supported EPA's discussion of validation in the guideline. Many of the participants seemed to believe that validation or calibration would somehow reduce the calculated concentrations and wanted to be able to use adjusted models in their final predictions. In this general context, some participants asked EPA to indicate performance criteria for various models.

A related topic was confidence intervals. It was recognized that any validation or calibration study would show a considerable scatter of the predictions around the best fit, and thus the question of confidence limits was appropriate. Several felt that the guideline should specify how one should make use of the confidence interval information in the interpretation of model results for comparison against standards. Some commentors requested that EPA publish in the guideline the confidence intervals for those EPA models which have already been validated.

A number of participants commenting on screening models supported their use. However, two commentors suggested that, wherever possible, more refined models should be used in order that consumption of PSD increment not be exaggerated by the use of such conservative models. A range of opinions was expressed as to the appropriate degree of conservatism in screening models.

More than 25 commentors suggested that the dispersion coefficients suggested for use in the models appearing in appendix B of the guideline are inappropriate for one reason or another. The first problem area concerns sampling time. It was generally stated that the sigma values in the models listed in appendix B of the guideline refer to three minute to ten minute averaging times, yet in the guideline, they are treated as if they apply to hour-long releases. Second, a number of the participants felt that the Pasquill-Gifford dispersion coefficients are not appropriate for use for tall stacks having initially large buoyant plume rise. A need for enhancement of both vertical and lateral dispersion coefficients in rough terrain was noted. One participant submitted rather extensive comments suggesting that the Pasquill-Gifford type A curve did not represent dispersion from tall stacks. Some commentors requested guidance on substituting other parameters such as delta T in the determination stability classes. One wanted to go to some method of continuous change of dispersion coefficients instead of resorting to discrete classes.

Another issue related to discrete classes of dispersion coefficients concerned the definition of urban versus rural conditions. It was suggested that the guideline should state criteria for choosing between the urban dispersion coefficient set and the rural set. Apparently because they believed that the Pasquill-Gifford dispersion coefficients used in the models recommended in appendix B do not apply universally, several participants requested that EPA provide guidance on how to adjust dispersion coefficients for the various factors that have been cited, such as rougher terrain and taller stacks.

More than 10 comments were received concerning the increments of modeling significance presented in Appendix A of the draft guideline. Some questioned whether there should be any such significance increments. Others supported the concept, stating that the increments should apply equally to PSD determinations in attainment areas and to nonattainment areas. There were requests for EPA to further explain its rationale in choosing the increments, and there were claims that the increments presented did not bear a reasonable relationship to the accuracy of monitoring.

Comments were received on the topic of receptor placements. The underlying issue is the spatial representativeness of the models. How closely must receptors be placed in searching for the maximum concentrations? The Texas Air Control Board made an explicit recommendation of one fourth kilometer to one-half kilometer for grid spacing.

In the San Francisco meeting some participants suggested that EPA should have to give specific justification for the recommended models. These commentors were looking for technical justification. For them, the fact that EPA had made Gaussian models widely available through the UNAMAP program, was not valid justification for recommending these models. The same commentors also requested that EPA provide more documentation on the models that are recommended in the guideline.

Several commentors from state agencies in the East requested more specific guidance on how to define nonattainment areas. Representatives of two or three state agencies suggested that sources of a size smaller than a hundred tons per year should be explicitedly included in the modeling, but it was recognized that one needs to establish a cutoff somewhere to keep the total number of sources manageable.

Five commentors objected to the requirement on page 29 and 30 of the second draft for some of the detailed power plant information called for. The argument was that the information called for had nothing to do with modeling and therefore was not appropriate to appear in a modeling guideline, although it might be appropriate elsewhere. Several participants in the meetings proposed that measured ambient air quality data should be used in preference to model estimates for air quality impact evaluations. It was implied that monitoring data were not currently used to the extend possible and relegated to a secondary role by the guidelines. This is not a modeling issue per se but, it has to do with the use made of modeling information in overall decision making. Although it was not explicitedly stated, the concern seemed to be over the degree of emission control required for existing sources rather than projected impact of new sources. In a related comment, one participant gave the opinion that a modeled "hot spot" should not be regarded as an air quality violation in need of further emission control, unless the violation had been confirmed by monitoring.

Two commentors suggested that EPA should somewhere define the division responsibility for carrying out modeling in new source reviews. These two commentors preferred that modeling be done by agencies rather than by applicants. Finally, a small group of commentors suggested that EPA needed to do more in the way of providing training for personnel in smaller agencies and in providing computer resources.

Appendix B

THE AIR QUALITY MODELING CONFERENCE
OF DECEMBER 14 AND 15, 1977

by

Gordon A. Beals, Ph.D
Member, AMS Committee on the
Meteorological Aspects of Air Pollution

Prepared for the Steering Committee
AMS-EPA Cooperative Agreement

December, 1980

THE AIR QUALITY MODELING CONFERENCE OF DECEMBER 14 AND 15, 1977

The conference was held pursuant to the new section 320 of the Clean Air Act Amendments of 1977. In the meeting notice, EPA announced its intention to focus on the EPA guideline on air quality models and that special attention would be given to modeling issues arising from EPA's program for preventing significant air quality deterioration. It was announced that comments received during this conference could also be considered part of the record with regard to EPA's proposed regulations to implement the P.S.D. provisions of the CAA Amendments. More than 300 individuals attended the conference, of which 50-60 were identified in the transcript. Over 40 written comments were also submitted.

A few of the comments were directed more at the framework of EPA's PSD program and role of modeling therein than at the techniques of modeling discussed in the guideline. The theme of all of this group of comments was that modeling alone was not suitable for evaluating the potential for significant air quality degradation. Another 23 comments were received repeating the issues raised in the earlier public meetings on the guidelines as concerns the relative importance that should be given to modeling as opposed to monitoring. Both industrialists and environmentalists seemed to prefer that the final determination of PSD consumption be based on monitoring rather than modeling. A compromise was suggested as follows: if EPA insisted on using modeling to keep track of PSD increment consumption, then the final estimate of increment consumption should be based on modeling that was calibrated to the sources in operation. Another four commentors raised the issue of conservatism as a policy. They generally stated that Congress had intended that modeling give a "best estimate" answer, not a conservative answer, when being applied to the PSD review process. Two more commentors raised a point which perhaps need not appear in the guideline, although these commentors felt it should. That point being that, with respect to short term PSD increment consumption, the consumption resulting from the operation of two separated sources will not necessarily overlap. One commentor felt that a radius of insignificant interaction should be defined such that outside the radius it would never be necessary to do simultaneous modeling of two or more source.

Specificity Issue

A major issue at the modeling conference as well as at the earlier public meetings concerned the degree of specificity versus the degree of flexibility suggested by the guideline. Thirty comments were received in this category. Ten of them directly supported the position that the guideline should be as flexible as possible. Another three took the position that a certain amount of standardization is desirable while maintaining the option of site-specific evaluation. One of these pointed out that, not only is it desirable to have consistency in the selection of models, but it is equally important to have consistency in the use of, and selection of input data, such as meteorological data and load levels. Another fourteen commentors supported the concept of case-by-case modeling determinations. Some of them suggested, as had been suggested earlier in the five public meetings, that the guideline should stress the desirability of an early conference between various parties to a modeling decision. One suggested that the EPA group at Research Triangle Park should develop guidance on accepting alternative models. This commentor felt that the Research Triangle Park group should also develop standard data sets to be used as inputs into the models to help the EPA Regional Administrators decide whether to accept an alternative model. Six commentors in this group also suggested that there should be some mechanism for an outside review of these case-by-case modeling decisions if any parties were dissatisfied with the outcome.

Four commentors brought up the newly published Workbook for the Comparison of Air Quality Models. Two of them felt that the workbook framework for comparing air quality models was conceptually wrong and suggested their own criteria. The suggested criteria were: (1) the comparisons should be based on accuracy not the degree of detail, (2) acceptable models should be consistent with physical reality, (3) that no model would be acceptable if it led to substantial underpredictions or overpredictions, and that a model should not be recommended unless it is supported by measurements. The other commentors, while not criticizing the workbook framework, felt that the inclusion of so-called reference models was a bad idea because these reference models might be mistaken for required models. They also complained that the workbook could lead to the incorrect impression that application of such a simple scheme by an inexperienced person could lead to a valid comparison of models.

Another large related group of comments discussed validation and calibration of models and model accuracy in general. The general points made were: (1) that validation or calibration should be done whenever possible, (2) that the present state of validation of EPA's recommended models is far from satisfactory, and (3) that high priority should be given by EPA to additional validation studies. Nine different people suggested that the guideline should contain specific procedures for conducting model validations, including, if possible, standards of performance. Some participants also asked that the material in the appendices to the modeling guideline be expanded to indicate the validation studies which have been performed on various models, and the range of conditions for which those validations were carried out. In validation, two types of information can be developed. The first is some sort of statistical measure of the overall accuracy and precision of the models. The second is information as to the sensitivity of the model to variations in the meteorological and load inputs. The participants felt that for the recommended models, both types of information should be presented in the guideline.

Assuming one does know the size of the error bars on a given model prediction and something about the sensitivity of the model to variations in inputs, one then needs some guidance on how to interpret this information. Six commentors specifically requested that the guideline discuss the interpretation of model uncertainties. Finally, six commentors suggested that, whenever possible, the model results should be adjusted to agree with the results of measured data. One of these commentors suggested that the modeling guidelines should also include specifications for how to use site specific model adjustments.

The Use of Meteorological Data

In the Washington DC Conference, participants either did not mention the length of data requirement or generally agreed with the guideline preference for multiple years of data. One commentor repeated the position stated during one of the five earlier meetings that the period of record should be thirty or forty years, corresponding to a lifetime of typical power plant. One participant discussed the interpretation of the results of multiple years of modeling. For purposes of comparison against the short term prevention of significant deterioration increments, this commentor said that Congress intended that the PSD increment might be exceeded once in a typical year, not that there should never be more than one occurrence per year in excess of the allowed increment. In line with this viewpoint, he stated that modeling additional years should serve to better define the typical year, not to search out extreme years.

Two participants mentioned problems with the data sets available from the National Climatic Center. The first of these had to do with NCC's use of three-hourly observations in place of hourly observations in its more recent STAR tabulations. This commentor wanted EPA to try to exert some influence with NCC to return to the processing of hourly data. The second commentor pointed out that in the NCC STAR tabulations, the frequencies are based toward certain wind directions when the data being processed comes from years after 1964. Under certain circumstances, this error can give rise to concentrations predictions 33% too high for certain wind directions.

Several participants requested that the guideline spell out requirements for on-site meteorological data. It was felt that the guideline should prescribe what meteorology to measure, what instrumental principle and instrument accuracy should be required, what length of record should be gathered, and what method should be used for stability class determinations for the on-site data. One commentor suggested EPA should include guidance on what kind of meteorological measurements should be used in evaluating releases from very tall stacks.

Two commentors suggested that the National Weather Service Airport data are not satisfactory with respect to representing direction. The main concern of these participants was that the National Weather Service observation uses a five-minute average wind direction near the end of an hour, not the average wind direction for the entire hour. These commentors would prefer that full-hour averaged winds be used for modeling, or else that the meteorological period be broken up into intervals shorter than an hour with a separate wind direction assigned to each interval. One commentor also suggested that the NCC method of determining the afternoon mixing depth had some well know difficulties. This commentor thought the guideline should indicate some sort of procedure substituting a default value when NCC value seemed obviously unreasonable.

With respect to the meteorological data requirements for modeling long range transport and complex terrain, one commentor pointed out that single station meteorology is not sufficient for these calculations, and that the type of data needed depended on the source-receptor distance. Another commentor suggested that there should be some mention in the guideline of how to use synoptic scale or mesoscale wind flow effects in complex terrain modeling.

There were two more repetitions of comments which had been made during the earlier public meetings about the treatment of calms in the CRSTER-RAM meteorological preprocessor. One commentor suggested using a random variation over a forty-five degree sector in place of the present practice of using persistent wind direction. There were also two comments concerning the assumed vertical temperature gradient used for calculating plume rise in stable conditions in the CRSTER model. One suggested the use of less steep temperature gradients when the stacks are quite tall. The other commentor said that there is too much difference between the assumed temperature gradient for stability category E and the stability category F, and that this jump produced too large a step in plume rise from one stability calss to the other. He suggested that some more refined measure of vertical temperature gradient was needed for the plume rise calculation subprogram.

Emission Rates and Operating Load

There was a limited discussion of this issue at the Washington meeting. Two commentors simply said that the assumption of a 100% operating load should not necessarily be required in all analyses. One person said that all EPA models of the sequential type should provide for hourly changes in emission rate. Two utility spokesmen made generally similar suggestions along the line of recommending use of load levels that can be reasonably expected to occur during each averaging time of concern. Finally, one again complained about the data requirement list on page 28 and 29 of the guideline, pointing out that some of the required data have nothing to do directly with modeling.

Dispersion Coefficients

The use of Pasquill-Gifford stability curves to describe diffusion from tall stacks was challenged from both ends of the spectrum. Two commentors suggested that vertical dispersion rates as fast as described by the Pasquill stability A curve do not occur in connection with tall stacks, while one commentor suggested that dispersion rates as small as stability F do not occur with tall stacks. One of these commentors and three additional participants recommended using the Brookhaven (ASME) dispersion curves in place of the Pasquill-Gifford curves, as recommended by the American Meteorological Society. One participant said there was serious controversy over the system used to represent atmospheric stability in the EPA models. The EPA method is based on low level meteorological data and no prevision is made for the use of more appropriate data even when it is available. Three participants discussed the possible enhancement of horizontal dispersion curves for periods longer than ten minutes. They requested a model adjustment so that the model would more nearly correspond to the expected horizontal dispersion in a full hour period. Another participant suggested that the models should provide for initial plume enlargement due to buoyancy. Finally, three commentors pointed out in need for the guideline to discuss the selection of urban dispersion coefficients versus rural dispersion coefficients. Some of these same participants also described their experiences in attempting to validate the RAM model using the urban dispersion coefficients. In general, they reported predicted concentrations exceeding observed concentrations by a large margin.

Screening Techniques

More than thirty comments were made concerning EPA's proposals to use the Guideline for Air Quality Maintenance Analyses Volume 10 screening procedures for preliminary review under the prevention of significant deterioration provisions. Almost half of these comments concerned the VALLEY model. Most of the commentors supported the screening model concept. Two or three opposed it on the ground that small sources approved on the basis of a screening model demonstration would be listed as having consumed so much of the PSD increment that there would be little room left for larger sources. Two commentors pointed out a way around this dilemma, which would be to approve small sources on the basis of a screening model demonstration, but then when modeling is done for larger sources, these previously approved smaller sources could be subjected to more refined analysis. In any case, they said, only a refined analysis should be regarded as the final estimate of PSD increment consumption for sources initially approved on the basis of screening model demonstrations. One participant claimed that dispersion from power plants was universal enough that a single set of worst case dilution factors could be applied for initial screening of all power plants.

On the issue of conservatism in screening models, there was a predictable split of opinion. Nine participants representing energy and industrial interests suggested that the screening models should not be too conservative. One participant representing environmental interests suggested that EPA should adopt very conservative screening models.

One participant pointed out that there was an error in one of the figures of the Guideline for Air Quality Maintenance Analysis.

The comments on the VALLEY model suggested either that this model should be deleted from all the guidelines including the screening guideline, or else that less conservative meteorological assumptions be made in applying this model.

Background

Six participants in the Washington Conference commented on background issues. The comments were generally similar to those made earlier in the five public meetings on the modeling guideline. In regard to provisions for growth, one of the commentors suggested that a ten year growth projection should become the standard because it was very difficult to plan for any longer period. Another commentor suggested that the wording should not be limited to emission growth but to background emission changes so as to provide for the possibility of long term emission reductions rather than emission increases associated with growth.

Guideline Review and Revision

Seven participants commented on revision of the guideline. Two of them suggested a need for a well organized, step-wise procedure for guideline revisions. One of these suggested a scientific advisory panel to review and evaluate proposed models. Another of the participants, without mentioning the step-wise feature, also requested that a wide cross section of the scientific community be permitted to submit recommendations on model revisions. One suggested a fixed schedule for the revisions so that everyone who would have input would have a definite target data for submitting new material. A state agency requested exactly the opposite procedure, suggesting that guideline revision should be made only on the basis of specific technical developments in modeling or changes in standards or policy, and that the revision should follow a formal notice announcement procedure. However, two other state agencies requested very frequent updates, something on the order of bimonthly modeling bulletin. Finally, one participant suggested something which had also been mentioned in the earlier modeling meetings, that sources undergoing review should be "grandfathered" against model changes during the review process.

Modeling Nitrogen Oxides

Three of the participants suggested that the recommended first-cut approach, which assumes all nitrogen oxides emitted are converted to NO_2 by the time the plume has a significant ground level effect, is too conservative. Three others suggested that NO_2 should be modeled only as a photochemical pollutant.

Significance Increments

Five persons commented again on the significant concentration increments which were listed in Appendis A of the second draft guideline and the interim guideline. In general, these comments suggested that the significant increments were not adequately justified and that they should be proposed in some other document.

Plume Rise and Inversion Penetration

Some commentors objected to various features of the Briggs plume rise formula. One of the participants preferred the larger numerical factor in the stable plume rise equation which had been given in Briggs' earlier works. Another suggested that he found the Holland plume rise formula to work better than the Briggs formula. Four participants recommended adoption of partial plume penetration through inversions. The reasons were split two and two. Two participants stated that full reflection produced excessive concentration estimates, while the other two objected to the feature of most models which assumes a zero ground level concentration if the calculated plume height is anywhere above the inversion base.

Short Term Concentrations Estimates

Three commentors suggested that use of the highest short term estimate instead of the highest, second highest estimate would not be appropriate. Instead, they suggested that some other method, such as statistical, should be used to estimate the second highest concentration at a given receptor. However, when sequential meteorological data are not available and one is faced with resorting to synthetic estimated meteorological conditions, then one does not know that this synthetic estimated condition will not occur more than once. One of the state agencies implied that in this case the synthetic estimate must be used as if the highest and second highest concentrations were equal.

Good Engineering Practice Stacks

Some participants commented on the GEP stack issue because it had been mentioned in the EPA notice of the proposed PSD regulations.

Three of the participants suggested that modeling stacks with artificial heights was an absurd procedure and implied that EPA should find some other way of compling with Section 123 of The Clean Air Act Amendments. One requested that the guideline provide guidance on what kind of obstacles should be the basis for deciding GEP.

Special Topics

There were quite a number of requests for guidance on various special topics. The topics mentioned included fumigation, aerodynamic downwash, plume interaction in complex terrain, line source models, fugitive emission modeling, long range transport, cooling tower impacts, visibility, and multiple source modeling. The concern in the multiple source situations was that the sequential hourly model approach such as RAM was too costly for multi-source modeling. On the other hand, such techniques as Larsen's statistical technique were felt to apply only to urban situations where there was a proponderence of small sources with emission points near the ground.

Complex Terrain and Plume Clearance Over High Terrain

A number of participants noted that the guideline contains no specific models for complex terrain and urged EPA to get on with the research needed to develop such techniques. In the present situation, one is left with no modeling recource except admittedly conservative screening models for the complex terrain case. As for plume clearance over elevated terrain, all of the comments received indicated that the standard EPA models did not work well. Most of the commentors preferred the Egan half terrain height correction. One state agency requested that EPA study the various alternative terrain correction schemes and apply in all of their models the one that works best.

Long Range Transport

A number of participants decried the lack of guidance for the long range transport situation. Several suggested possible model modifications. One of the more common suggestions was that concentrations should be set to zero if the wind is not persistent enough to advect a puff from the source all the way to the supposed terrain impact point. Other suggestions included accounting for deposition, decay, and washout in sequential models if they are to be applied in the long range transport situation.

Other Miscellaneous Comments

A few participants requested guidance on modeling receptor locations. They asked that the guideline specify the minimum grid spacing. Others said the guideline should state that receptors should not be located on the source's property or beyond the prediction capability of models or in places generally not accessible to the public. A few commentors suggested that the Gaussian kernel does not represent state-of-the-art in all cases. For these cases, physical, numerical, or other methods might be available and superior. One person suggested that the APRAC 1-A model should be deleted and replaced in the guideline by APRAC 2. The thought was expressed that it was generally desirable that models should allow for as wide a range of meteorological and source conditions as possible. Others suggested that methods, rather than specific models, should be the focus of the guideline recommendations. One spokesman requested that when an environmental regulatory agency has modeled a source as part of the review process, then that agency should issue a comprehensive report documenting the results of the analyses.

Some general comments were made concerning the suitability of models for inclusion in the guideline. One was that the mere availability of models to state agencies is not an appropriate criterion for model selection. One utility stated that proprietary models should be accepted when necessary, while others thought that proprietary models should be rigerously excluded.

There was a challenge to a statement on page 11 of the interim guideline which implied that a 15 ton per year source could produce concentrations exceeding the proposed significance increments even with a stack height greater than 30 meters. Other isolated comments included (1) that there was a need for a critical review of photochemical models, (2) that the EPA emission offset policy is scientifically unsound, and (3) that in the CRSTER model the separation between stacks should be accounted for. There was a suggestion that something called generalized persistence technique be included in the guideline as one possible method of analysis, and a statement that model building is a form of science and should be divorced from regulatory agencies.

Finally, there were several comments on the EPA RAM model. Several of these reported unfavorable experiences, such as predictions of concentrations considerably higher than measured. Others simply complained about the expense of running such a model, and one person called attention to the fact that there still was not a RAM users guide as of the time of the Washington conference.

However, one local air pollution control agency went so far as to state that they preferred using RAM for annual averages in place of the more commonly used sector average models, such as CDM or AQDM.

In summary, the comments in the Washington DC modeling conference were generally similar to the comments which had been presented earlier in the five public meetings on the use of air quality models. Some new topics were raised because the Clean Air Act Amendments of 1977 were passed in the interval between the first five public meetings and the December modeling conference.

Appendix C

COMMENTS ON THE ADEQUACY OF E.P.A. RESPONSES TO THE ISSUES RAISED IN THE PUBLIC MEETINGS AND AIR QUALITY MODELING CONFERENCE

by

Gordon A. Beals, Ph.D
Member, AMS Committee on the
Meteorological Aspects of Air Pollution

Prepared for the Steering Committee
AMS-EPA Cooperative Agreement

December, 1980

COMMENTS ON THE ADEQUACY OF E.P.A. RESPONSES TO THE ISSUES RAISED IN THE PUBLIC MEETINGS AND AIR QUALITY MODELING CONFERENCE

1. INTRODUCTION

In addition to a summary of the comments of the participants in the five public meetings and the Modeling conference, the AMS technical review committee requested an opinion on the adequacy of EPA's responses to the issues raised in those meetings. Such a judgement is naturally dependent on the perceptions of the reviewer. The comments which follow are not intended to assess whether EPA has any legal obligation to respond or whether the agency's responses (if required) are legally adequate. The comments are meant to indicate whether EPA's actions and explanations are technically satisfying.

Many of the broader issues are discussed or alluded to in the Introduction (Section 1.0) or the Overview (Section 2.0) of the guideline. These broader issues will be discussed first, followed by some specific comparisons between the April 1978 (final) Guideline on Air Quality Models and the second draft (the May 1977 version of the guideline.)

2. DISCUSSION OF ISSUED RAISED IN THE INTRODUCTION AND OVERVIEW

2.1 Purpose of the Guideline

In the first paragraph, the guideline says "The guideline identfies those techniques that EPA considers acceptable." By inference, this means that all other techniques are not acceptable. The majority of the modeling community felt that there are other models and techniques, besides those specifically recommended in the guideline, that should be found acceptable. Thus the majority would feel that the guideline has not yet come up with a common measure of acceptable technical analysis.

The second paragraph of the Introduction states that the preferred model is that which best simulates atmospheric transport and dispersion in the area of interest. An overwhelming majority of the participants in the modeling conferences would agree with that statement. Numerous conference participants felt that the guideline is deficient because it fails to provide common measures of performance by which it can be judged how well a model simulates transport and dispersion in a given area of interest.

Many aspects of the guideline seem aimed at defeating a search for the best simulation of transport and dispersion in an area of interest. Ardor for the search is dampened by two other sentences in the second paragraph of the Introduction: "It [the guideline] should be followed in all evaluations relative to State Implementation Plans (SIPs)" and "However, deviations from this guide should be fully supported and documented." It is as if E.P.A. feels that striving for the best representation of transport and dispersion in individual situations is not consistent with "consistency." This conflict will be discussed further in the next section.

Another purpose of the guideline is to specify data bases (meteorology, background air quality) needed for an adequate technical analysis. It is felt that the guideline has been more successful and more responsive to inputs from the modeling community with respect to this purpose than with respect to the recommended models.

In one respect, the guideline has almost the force of a regulation. The guideline is used as the vehicle by which E.P.A. "specifies with reasonable peculiarity" the models to be used in new source reviews under the Preventions of Significant Deterioration (PSD) section of the Clean Air Act Amendments of 1977. In its PSD regulation, EPA requires use of the procedures, data bases and models specified in the guideline unless a wavier is obtained from the Administrator.

The PSD regulations, which were open for discussion at the Washington DC conference, specify two other documents, the Guideline for Air Quality Maintenance Planning (Vol 10) and the Workbook for Comparison of Air Quality Models. Some comments on these publications were received at the Washington conference. Any overall commentary on the adequacey of EPA responses to modeling comments must also include some mention of these publications.

2.2 Consistency

Throughout the history of air pollution control there have been two methods of approach, one emphasizing uniform emission limits while the other emphasized uniform air quality standards. Herschel Slater mentioned these approaches in his opening remarks at all of the air quality modeling meetings. Most meteorologists believe that there is not a single relationship between emission rates and air pollutant concentrations for various downwind distance that would apply to all sites. Therefore, it is not possible to consistently emphasize both the emission limitation approach and the air quality management approach. Complete uniformity of emission limitations between two regions having identical anthropogenic activity would necessarily result in unequal concentrations of air pollutants. On the other hand, rigid adherence to the principle of air quality management resulting in equal maximum air pollutant concentrations would necessarily require unequal emission limits.

In implementing the uniform emission limits approach, one needs only a simple model which will specify approximately the maximum concentrations versus emissions and stack height in terrain of "average" roughness and average cloudiness. But to fully implement the air quality management approach, an infinite number of models is required. Each site will have its own unique emission-concentration relationships. Consistency in this approach means using a consistent scheme to parameterize the characteristics of atmospheric turbulence as it varies from site to site.

The choice of approach must be decided by E.P.A. and Congress. Meteorologists can only advise that optimum implementation of the air quality management approach requires models with complex representation of transport and dispersion. However most meteorologists at the modeling conferences were convinced that the air quality management approach is practicable and that it is the accepted approach in the United States. For them, consistency would be served by having a uniform attitude toward site-specific adjustment of model parameters and uniform criteria for judging the goodness of such adjusted models.

2.3 The Degree of Flexibility vs the Degree of Specificy

To serve as a guide the document must be reasonably specific. It should foster more uniform application of models so that similar results are obtained by different well-trained analysts. Such a similarity of results is desirable so that control agencies at various levels of government make consistent control decisions within and among their organizations, so that applicants for permits understand the citeria and interpretations that are used in analyzing the impact that the proposed facility makes on air quality, and so that uncertainties in the reviewing and permitting process are reduced. There is a clear need for air acceptable analytical approach by which an applicant can test the likelihood of compliance before significant investments are made in the pre-construction phases of new development. I don't believe there is the slightest disagreement within the modeling community about the desirability of such consistency.

The disagreement, a serious one, concerns the appropriate way to achieve the desired similarity of results. One way is for all groups to use a common set of models. The question then is whether the models listed in the guideline give a sufficiently good simlation of the transport and diffusion at a wide variety of sites that they are universally suitable for implementing the air quality management approach described above. The answer from the private sector and some state-level agencies was then a resounding No! Some participants claimed better agreement in post-operational monitoring when using other models. Others argued, from a more theoretical viewpoint, that most situations are a combination of circumstances which cause each site to be unique. Still others claimed that the peak concentrations predicted by the EPA models were simply too high. Subsequent validation studies showed that the EPA models give reasonable results in several typical situations. However, other studies in other places showed that other models such as those based on the Brookhaven dispersion equations frequently perform better than the EPA models.

One of the main areas where conference participants felt that EPA models could be improved upon was the domain of dispersion coefficients. A number of participants sought the flexibility to depart from the Pasquill-Gifford coefficients where such a departure could be justified. The participants did not say how they would modify the dispersion coefficients. Presumably, most would increase the dispersion rates (except for class A) at most sites because the Pasquill-Gifford curves were developed from experiments in very flat terrain. These participants would be dissatisfied that the guideline still does not contain any explicit discussion of dispersion coefficients.

An alternative to the use of standard models, suggested by several participants, was for the guideline to include standards for model performance and criteria for the representativeness of the input data and for validation procedures. Any model which met the standards should produce results sufficiently close to other models meeting the standards that different agencies or interests would make similar decisions. A major obstacle to implementing this type of approach is that appropriate model performance criteria have not been devised. The participants who favored this approach were of course disappointed that such criteria had not been proposed. However it should be noted that E.P.A. has been funding research aimed at evaluating model performance criteria.

Some of the participants who favored a strictly case-by-case approach for modeling specific sites also sought a mechanism for review of a Regional Administrator's decision. Such review would be at E.P.A. Headquarters or even through "peer review" outside of E.P.A. The agency has not publicly indicated its response to this proposal.

It seems that the main thrust of EPA's response to the flexibility issue has been to bring forth the Workbook for Comparison of Air Quality Models. Through this workbook, the EPA seems to have established a mechanism for granting all the flexibility that anyone would want in adopting alternative air quality models. The real measure of EPA flexibility, is found in how readily EPA agrees that an alternative model may be better than one of the recommended models. My impression is that at least in the eastern half of the country EPA Regional Administrators have been very reluctant to accept the viewpoint that alternative models may be better than the EPA models without referring the matter to the staff at Research Triangle Park.

2.4 The Use of Models to Supplement or Replace Measured Air Quality.

Several participants at the five public meetings and the Washington DC modeling conference proposed that air quality data should be used in preference to model estimates for air quality impact evaluations. At the modeling conference the comments were mostly triggered by E.P.A.'s proposal to use modeling exclusively to track "consumption of PSD increment."

The guideline now states in the final paragraph on page 4 that, whenever possible measured air quality data should be used to determine the accuracy of model estimates, and that information on the accuracy of a model should be available prior to the use of that model to evaluate control strategies or to determine allowable emissions. This emphases on measured data and model performance is responsive to inputs received during the public meetings and modeling conference.

Page 5 of the modeling guideline states that EPA has never encouraged the use of air quality models in place of measured data. Several participants would find this statement irreconcilable with EPA's position on the use of models to track PSD increment consumption, and with the statement on page 9 of the guideline referring to modeled concentration estimates as providing "a more conservative and realistic indication of the potential for NAAQS violations and of the appropriate emission limits than do actual measurements at few monitoring sites." Another passage which might be construed conflicting with the overview is the statement on page 34 of the guideline, which says "However, where the source is not isolated, it may be necessary to use a multi-source model to establish the impact of all other merely sources during dispersion conditions conducive to high concentrations."

The sore point with some industrial representatives was probably the "hot spot" issue, although it was seldom mentioned directly. By "hot spot" I mean a modeling receptor that indicates a violation of air quality standards in an area where all the monitored data show compliance. One participant zeroed in on the point and said that EPA should not regard such a "hot spot" as an actual air quality standards violation unless the violation had subsequently been confirmed by monitoring. Another said that the answer to insufficient monitoring should be more monitoring, not modeling. The technical issue here is whether models are accurate enough to be accepted as evidence of a probable air quality standards violation.

About a dozen comments favored the use of calibrated models together with monitoring data to define the extent of problem areas. Another ten opposed either EPA's proposal on PSD increment consumption or the use of modeling to reject SIP's, or both. The PSD increment consumption issue has by now been litigated in the courts so continued argument of it would be inappropriate.

2.5 Screening Techniques

In general, EPA has been responsive to the suggestions of the modeling community with regard to the incorporation of screening techniques. However, some of the participants in the Washington DC modeling conference believed that EPA had been too conservative in the choice of meteorological inputs to the screening techniques. For example, the screening guide suggests making a calculation of stable plume rise and then cutting the plume estimate rise in half. But because of the cube root dependence of the stable plume rise formula, this really amounts to making an eightfold change in any one of the input parameters. No one would object to EPA's claim that the procedures have been chosen from public materials which have been subjected to scientific review for several years. However, this claim does not extend to the recommended input meteorological values.

2.6 Guideline Review and Revision

The participants in the five public meetings and the Washington DC modeling conference unanimously recommended review of the guideline at least annually. Many participants, however, made a distinction between review and revisions, the latter to occur only when warranted by scientific developments. It was left unsaid whether revisions would be made on the basis of major or minor scientific developments.

Several participants expressed concern over the possible impact of a revised guideline on control strategy developments and new source reviews in progress. The importance of continuity was brought out and some participants suggested fixed intervals for revisions, such as 3 to 5 years. It was felt that the same model should be used throughout the review period of a specific project, and there should not be retroactive evaluations based on new models introduced by revisions of the guideline. However, it was recognized that revisions might be required to maintain consistency between the guideline and regulations that might be changed in response to legislation.

The EPA response to comments, as represented by the content of the last paragraph in the Preface of the final guideline, seems to be adequate.

3. COMMENTS CONCERNING SPECIFIC SECTIONS OF THE GUIDELINES

Section 3

The second paragraph on page 9 remains unclear. If the source in question is known to have little impact, then why is this source being modeled for this monitor location? The sentence which reads "For the purpose of specific emission limits these measured concentrations should be ranked ahead of the estimated concentrations..." is unclear. EPA needs to elaborate this paragraph somewhat more fully.

The final paragraph of section 3.1 of the guideline seems to reiterate without change the position of EPA recommending in some circumstances the use of the highest short term concentration instead of the highest second highest concentration estimate. EPA has not really explained why they continue to hold this position. It may be that EPA is responding to a concern, expressed by at least one of the state agencies, that if a concentration is found once in a year, then how can one be sure that an identical modeled concentration would not occur a second time? Several participants, felt that, instead of requiring use of the highest concentration automatically, statistical techniques should be used for establishing the probable second highest. It could be that applying statistical techniques to the set of modeled concentrations would show that the highest modeled concentration could be expected even more often than twice a year. EPA needs to be more explicit about the circumstances in which use of the highest modeled concentration is appropriate.

The first paragraph of Section 3.2 is considered to be a change which is responsive to the comments of the modeling community. The significant modeling increments have been removed from the guidelines and now appear in other regulations.

The final sentence of Section 3.2.1 refers back to the final paragraph of Section 3.1 regarding use of the highest short term concentration. The comment given above on the final paragraph of Section 3.1 applies here as well.

Section 4

At several points, throughout Section 4, EPA refers to the guideline models as state-of-the-art. A few of the participants strongly disagreed that the Gaussian approach is state-of-the-art. There apparently is not a common understanding of the phrase "state-of-the-art." In the sense that EPA uses the phrase here, it is very much like saying that the internal combustion engine is state-of-the-art in automobiles. After all, the Gaussian solution was published by O.F.T. Robert in 1923. The application of Gaussian modeling to elevated industrial sources goes back at least to 1947. Several authorities, including Monin in Russia and F.B. Smith in England, now state that the vertical profile of pollutant concentration from surface releases, in fact, is not Gaussian. At almost every scientific meeting devoted to air pollution modeling in the last five years there has been at least one paper describing some numerical model which is shown to give a better representation of the vertical profile of pollutant concentrations than an appropriate Gaussian model. In the straw vote, which was taken at the Salt Lake City Conference a majority of those conferees who considered themselves as having a primarily academic affilation voted in favor of non-Gaussian models for complex terrain applications.* This group, along with most modelers from California would not agree with characterizing Gaussian models as state-of-the-art. I would have been happier if EPA had simply said Gaussian models are recommended without trying to constantly characterize them as state-of-the-art.

It is felt that the changes on the top of page 14 have been responsive. Likewise, EPA's position on photochemical oxidant modeling is responsive to the concerns expressed in the various meetings. However, the participants in those meetings, I am sure, felt that it would be preferable if EPA could have had an acceptable technique ready to recommend by now.

---

*During the panel discussion period following the session on complex terrain at the Salt Lake City AMS/APCA Joint Conference on Applications of Meteorology to Air Pollution, Don Henderson, the EPA Region VIII Regional Meterologist, took a poll of the audience concerning the participants' preference for Gaussian or numerical models for a relatively simple situation in complex terrain. The application was the calculation of maximum SO_2 or particulate concentrations 25 km downwind of a large source located in complex terrain. Mr. Henderson asked the members of the audience to categorize themselves as belonging to: (a) academic or laboratory, (b) regulatory, (c) industrial or (d) consultant or other affiliation, and to vote by category of affiliation on the question of whether to use a Gaussian or numerical model for this application. The group considering themselves as having primarily academic or laboratory affiliation voted narrowly in favor of using a numerical model. All the other groups voted heavily in favor of the Gaussian model. Two conclusions can be drawn from the vote. One is that, most users still preferred the Gaussian model. The other is that a significant sub-group, consisting of the people most active in pushing forward the state-of-the-art, did not consider the Gaussian model to be adequate. Even among the participants who voted in favor of the Gaussian model for this application, some apparently voted on the basis of economic instead of scientific considerations. One consultant turned to me after the vote and commented that, regardless of the merits of numerical modeling he would not ask a client to bear the additional cost of numerical modeling for the application proposed by Don Henderson.

A similar comment applies to the second paragraph on page 15 of the guideline. EPA's refusal to recommend any single model for complex topography or marine influences is responsive to the inputs by the modeling community. However, some participants lamented the absence, at that point in time, of any techniques that could be recommended.

The whole of the introductory two pargarphs of section 4.3 seem to be an improvement except for the just mentioned characterization of Gaussian models. On page 20, the first full paragraph states that where a refined analysis is needed and there are no significant meteorological terrain complexities, the single source (CRSTER) model is recommended for use. At the end of the paragraph, EPA, in refering to alternative models, says that where these models estimate essentially the same concentration as the EPA model, modified computer codes are acceptable. What this paragraph does not do is indicate what other models might also be acceptable. Such an indication of acceptable alternative models was sought by a number of participants in the various modeling conferences. What the modeling community is seeking is a recognition on the part of EPA that some models which give results different from the EPA models could be just as acceptable as the recommended models, if not more so.

There are many reasons why a valid model may be a good representation of the transport and dispersion in a particular area and yet produce concentration estimates differing by a factor of three or more from the EPA model at some receptors. For the most part such differences would be the result of using different dispersion coefficients or different treatment of terrain. A numerical model would also be expected to differ from the EPA model at some receptors. The modeling community is anxiously awaiting the day when the acceptability of alternative models can be judged on the basis of objective performance criteria.

The paragraph at the bottom of page 20 is certainly responsive to comments received during the various meetings. Again the modeling community regrets that up until now there has not been enough research devoted to complex terrain modeling for EPA to be able to make a satisfactory recommendation.

Section 4.3.2 and 4.3.3 continue the trend toward minimal use of the rollback model. Here EPA has been responsive to the recommendations received in the various public meetings.

Another improved section is section 4.3.4, models for carbon monoxide. On the other hand, section 4.3.5 dealing with models for nitrogen dioxide is not felt to be fully responsive. The conference participants did not feel that rollback was a viable option for reactive pollutants.

Section 4.4 of the guideline is not appreciably changed from earlier versions. The conference participants will be pleased to know that EPA is conducting research aimed at providing the guidance which the conference participants sought in these special situations.

Section 5

Some participants in the modeling conferences objected in principle to including in a modeling guideline some of the source characteristics and operating conditions now listed on pages 28 and 29 of the final guideline. The point made by the participants is that many of the listed items, such as pollution control efficiencies, are not model inputs at all, but are means for checking the accuracy of model inputs. Who cares about the boiler layout as long as the stack emissions are correctly specified? If the guideline is not to be a checklist for new source review, then it was felt that only items directly usable in modeling, such as stack parameters and emission rates, should be listed. If the guideline is to be a checklist for new source reviews, then the listed items may be appropriate, but EPA should clarify which items are essential model inputs and which items are required as supporting information.

The changes in section 5.2 on meteorological data seem to have all been for the good. There are some definite improvements in the guideline. In particular, the second paragraph on page 31 of the guideline makes a definite improvement over earlier versions. Again, the material on page 32 of the final guideline represents a distinct improvement; however, the conference participants had requested guidance on specific checks for the representativeness of meteorological data for a particular site. Such guidance is not yet in the guideline.

The conference participants will be disappointed that Section 5.3 of the guideline still does not provide any guidance on the minimum acceptable receptor spacing.

Section 5.4, Background Air Quality, presents a definition of background air quality in response to comment made at the public meetings. The definition given, "background air quality relevent to a given source includes those pollutant concentrations due to natural sources and distant, unidentified man-made sources," responds to those particpants who wanted to use the word "background" the way it is used here and who wanted to substitute the word "baseline" for air quality effect of all sources not explicitly modeled. I feel that for most readers the change in the present guideline is a step backward from earlier versions because a majority of users accept the term "background" to mean the result of all sources, natural or man-made, that are not explicitly modeled.

Finally, as a minor point, EPA repeats on the bottom of page 36 that growth caused by nearby sources should be considered for the next ten to twenty year period. Most of the commentors preferred to cut if off at ten years, feeling that it was not possible to estimate growth any further ahead than ten years.

Section 6

EPA has been responsive on the issue of sensitivity tests. The inclusion of the concept of sensitivity tests in the paragraph at the top of page 41 of the final guideline very definatly has been responsive to the expressions of the modeling community. Unfortunately, it has not been possible for EPA to be as responsive on the issue of calibration of short term models.

The issue of validation was certainly widely spoken to in the five public meetings on the air quality modeling guideline. Many of the commentors spoke about the EPA models as being unvalidated. This now seems to have been at least partly the result of slow communications. At the time that the guidelines first appeared, the only validation of EPA models had been in EPA reports. Burt's paper on the validation of the VALLEY model might have been in press, but the Salt Lake City conference had not yet taken place when the majority of the comments about unvalidated models were made.

It is noted that the model summaries in Appendix A make some brief references to validations. However, quantitative results are presented only for AQDM. There are no citations of references where the actual results could be found. It may also be appropriate to point out that none of the EPA validation studies has ever been reported in a peer reviewed publication where the full manuscript was available to reviewers before acceptance.

Appendix D

SUMMARY REPORT ON EPA'S 1980 AIR QUALITY MODELING CONFERENCES

by

Gordon A. Beals, Ph.D
Member, AMS Committee on the
Meteorological Aspects of Air Pollution

Prepared for the Steering Committee
AMS-EPA Cooperative Agreement

June, 1981

SUMMARY REPORT ON EPA'S 1980 AIR QUALITY MODELING CONFERENCES

The Environmental Protection Agency held a series of three two-day public meetings in October, 1980 for the purpose of presenting the agency's proposed revision to its 1978 Guideline on Air Quality Models and receiving review and comment by the public on these proposed revisions. The proposed guideline revision was intended to be synchronized with the triennial conference on air quality modeling required by Section 320 of the Clean Air Act. The comments received during the October, 1980 meetings were to be weighed and a final guideline presented at the Section 320 conference in early 1981.

The guideline revisions had been made as a result of: (1) a series of annual internal EPA workshops, (2) a cooperative agreement with the American Meteorological Society, (3) a solicitation of new models from the technical and user community, and (4) on-going EPA research programs. The revisions presented in October 1980 were largely a result of the internal EPA workshops. The cooperative agreement with the American Meteorological Society produced some suggested criteria for objectively comparing one model to another. However, the criteria were made available only at the last minute and are not yet a part of the draft guideline. Similarly the results of the model solicitation were not incorporated into the October, 1980 draft, but EPA's tentative decisions regarding the submissions were announced at the Washington Conference. A few of the conference participants commented on these recommendations as well as on the proposed guideline revisions. One set of comments was keyed to a summary of the EPA internal workshops, rather than the proposed guideline revision.

The precise number of participants was not stated anywhere in the materials available for this summary. The transcripts contain formal statements by more than 40 individuals not associated with EPA, not including presentations at each meeting by members of the Steering Committee of the cooperative agreement with the AMS. The transcripts also reflect participation by five to ten other people at each meeting. More than 70 sets of detailed written comments were received.

The comments addressed a wide range of topics including eight on which EPA specifically solicited advice. The eight items on which EPA formally solicited advice and comment were:

1. Specific changes to EPA recommended models.

2. Screening and refined techniques for complex terrain.

3. Requirements for emissions, meteorology, and air quality data bases, including the determination of background.

4. Selection of receptor sites.

5. Criteria for the case by case acceptance of non-guideline models.

6. Use of measured data in lieu of model estimates for establishing emission limits.

7. The appropriate number and placement of monitors to support a model evaluation.

8. Performance measures and standards for models.

Most of the specific comments were critical of the proposed guidelines and represented in industrial viewpoints. Commenters gain nothing by stating agreement; thus they reserve their words for points of disagreement. The sections which follow summarize the comments that were made, focusing mainly on prepared statements and written comments. The paraphrasing of the comments doubtless reflects the reviewer's perception of the issues.

Recommended Changes to EPA Guideline Models

The first major topic on which EPA had requested guidance in these conferences was the question of what changes should be made to the EPA guideline models. There were approximately 250 comments in this area, counting both written comments and oral comments. Comments on the guideline models were received in the areas of plume rise, terrain clearance, dispersion coefficients and a considerable array of other topics.

Plume Rise

All of the EPA recommended models for point sources use Briggs' formulas for computing terminal plume rise. However, several commenters pointed out that the various models are not consistent with one another in regard to transitional plume rise. Some models use terminal plume rise initally while other models use transitional plume rise. Three commenters recommended using transitional plume rise in all the models, while three others preferred using the final plume rise in all the models. It was also noted that some of the models, RAM in particular, use stack tip downwash while many others do not. The three people who commented in this area preferred the use of stack tip downwash in all the models. One person suggested that all of the models should provide for momentum plume rise with a crossover to a bouancy dominated plume. Two others suggested that plume rise enhancement for sources in close proximity to each other should be incorporated in the models. A fourth person suggested that all of the models should use the same numerical constant for calculating plume rise in stable conditions.

Terrain Clearance

Closely related to plume rise is the treatment of undulating terrain. The EPA models recommended for single source rural applications reduce the effective plume height by the difference in terrain height between stack location (unless the receptor is lower than the stack base). One commenter noted that there were inconsistencies between CRSTER, ISC, and MPTER in regard to this terrain adjustment. Another commenter said that terrain corrections in all EPA models should be made more realistic. A third commenter recommended that the CRSTER model should be modified to provide for some lifting of the plume approaching higher terrain.

Wind Profile

All of the EPA recommended models for point sources use a power law expression to extrapolate wind speed from an assumed anemometer height to stack height. The power law exponents vary with stability class. Four comments noted that the sets of power law exponents differ from model to model. One person commented that the differences between the profiles in different models are soundly based. Another person suggested that all models use the most recent set of profile exponents. A third commenter recommended retaining the particular set of profile exponents that are used in the CRSTER model and the CDM model. The reasons for the recommendations are, first, that this set is based on published work by deMarrais, and secondly, that EPA had never published the bases of the other exponents. Two comments suggested that EPA models should provide for user input anemometer reference height, and two comments suggest providing for user input of the wind profile exponents for various stability classes. As a final comment in this area, one person claimed that the present procedures could overestimate the wind at the stack top in the case of tall stacks.

Dispersion Parameters

There were quite a number of comments on the dispersion coefficients which are used in the various guidelines models. Twelve commenters noted that the presummed urban dispersion coefficients, the coefficients of McElroy and Pooler, were not used in all the models that presumably can be applied to urban situations. In particular, the CDM and ISC models did not use the McElroy and Pooler coefficients in the urban situations. Four commenters proposed using the McElroy and Pooler coefficient in the CDM and ISC models except for receptors very near buildings where enhanced turbulence or downwash was modeled. An option to use buoyancy-enhanced initial dispersion exists in the most recent EPA models, but not in the earlier models. Seven participants noted this difference; four recommended that buoyancy-induced dispersion be included in the models.

Four commenters attacked the use of Pasquill-Gifford Class A for tall stacks in rural areas. Another comment said that cogent validation data are still lacking. Still another comment said that the CRSTER model over-predicts in Class A conditions and under-predicts the annual average. (Such behavior implies that the ration of sigma z to sigma y in the CRSTER model is too large in Class A conditions, but too small in the most frequent conditions).

Quite a number of comments suggested modifying at least sigma y in some fashion. Ten suggested explicitly that the Pasquill-Gifford sigmas should be adjusted for averaging time. One proposed dividing the hour into ten-to-fifteen-minute subperiods, provided continuous wind data are available. Three commenters specifically recommended that EPA adopt the conclusions of the Summer 1977 American Meteorological Society workshop. Another suggested that all of the models provide for use of split sigmas, with the bulk Richardson number as the criterion for sigma z. One other recommended terrain adjustments similar to those recommended in nuclear regulatory commission Reg. Guide 1.145. Two commenters simply requested that the guideline contain explicit guidance on the selection of dispersion coefficients. Two others questioned the use of dispersion coefficients based on tracer experiment conducted essentially at ground level to represent dispersion from tall stacks. One comment stated that enough was known about dispersion coefficients already to suggest that the Pasquaill-Gifford coefficients in the EPA models should be scrapped and replaced by the dispersion and coefficients of the ASME Guideline. Another commenter stated that EPA should conduct more research into atmospheric dispersion. One comment stated that a scheme should be developed that allows for continuous, rather than discrete, variation of dispersion parameters with stability. Finally, one person suggested that the work of Irwin at EPA should be adopted and the guideline models altered accordingly.

Mixing Heights

There were six comments on the way that mixing heights are assigned by the preprocessor subprograms. Two of these comments suggested that the mixing height should not be allowed to approach zero even in rural conditions. One suggested a minimum height of 50 to 70 meters. Two other comments suggested that temperature data should be used in the process of interpolating mixing heights from hour to hour. One of these people also stated that those surface temperatures involved in computing morning and afternoon mixing depth should be taken from the same site as the radiosonde launch. One apparently wanted to be able to substitute more frequent actual radiosonde data in place of the standard mixing heights interpolated from maximum and minimum afternoon or morning mixing heights. Another simply stated that theguideline should provide guidance on interpolation of missing mixing heights.

Another set of comments related to the interaction between plumes and inversions. One pointed out that the comparison of mixing height with plume height in the CRSTER model is made independently for each receptor point. He recommended that the comparison algorithm be changed so as to avoid a situation in which the plume is calculated to be above the inversion near the stack but below the inversion at some distant elevated terrain.

Other Recommended Changes

One comment pointed out inconsistences between the RAM model and the "urban mode" of the CRSTER model. This comment recommended deleting the urban mode of CRSTER. The comment continued, stating that the flat terrain assumption of RAM is inappropriate and should be modified, but with less terrain adjustment than CRSTER. One state agency suggested modifying CRSTER to allow the input of additional receptors not located on rings and to allow for meteorological input not generated by the EPA pre-processor. One comment endorsed the minor changes in CRSTER procedures recommended in the October 1980 proposed revision. Two comments suggested modifying all models to accept hourly variation of input emissions.

Additional Model Comments

There were a large number of comments on other aspects of guideline recommended models but which did not make any specific suggestion for changes. There were an even dozen comments concerning the use of Gaussian models at various distances downwind, and whether the Gaussina model itself should be considered "state-of-the-art". Four said specifically that the CRSTER model should not be used for distances more than 50 kilometers from the source. Another questioned the use of Gaussian models beyond 30 kilometers, while one commenter even questioned the use of Gaussian models beyond 10 kilometers. Continuing with this theme, one commenter suggested that Gaussian models should not be used beyond 50 kilometers in level terrain or beyond 10 kilometers in complex terrain. The same commenter stated that Gaussian models shoiuld not even be used for screening pruposes beyond 50 kilometers. As authority for this statement, he referred EPA to the paper given by Arthur Bass at the New Orleans AMS meeting.

There were five comments concerning either the ISC or the MPTER model. The sense of these comments was that EPA should clarify the status of these models (that is, whether or not they are recommended models within the meeting of the guideline), or else calling the attention to a need for further check out and clarification of exactly which version of the model is the official version.

There were two comments about line source models for mobile source applications. First the California DOT stated that they agreed with the application limits suggested by EPA for the CALINE 3 model. A spokesman for General Motors wondered why the HIWAY 2 model is listed by EPA in the guideline instead of the GM line source model. His several arguments included a statement that the GM line source model has been available in the open literature for the past two years while HIWAY 2 has not been published. He claimed that the GM line source model is not less refined than HIWAY 2, and that, according to a study by the New York State Department of Environmental Conservation, the GM line source model performs as well as even HIWAY 3 or 4. He further claimed that the GM model uses a theoretically superior line source algorithm. The GM model was claimed to be amenable to physical interpretation, and the correlations between the GM line source model and data was claimed to be as good or better than for HIWAY. This spokesman also noted that HIWAY 2 and CALINE 3 both contain options to predict concentrations for configurations for which there has been no validation.

There were several comments directed at the EPA document titled, "Summary of Proposed Recommendations Concerning Air Quality Models Submitted to the Environmental Protection Agency". In most cases these were comments by the various models' authors seeking to correct statement EPA had made in the publication. These comments will not be discussed further. One commenter stated that EPA should not give "recommended" status to the ERT long range transport models.

Several participants commented on inconsistencies in EPA recommended models for urban areas. In particular, the models RAM, CDMQC, and ISC, and some cases CRSTER, are all recommended for different types of applications in urban areas. However, there may be up to three different treatments of the dispersion coefficients. An example of the difficulty is that the RAM model is recommended for a short term applications, but if one uses the RAM model to calculate a series of 365 daily values, and averages these results, the answer will be quite different than would be predicted by the CDMQC model for annual average. There is also a subjective element in applying these models. Once commenter used the example of the Houston Texas ship canal and asked if it is an industrial complex or an urban area. One commenter suggested that the ISC short term model be used instead of the RAM model for urban multipoint short term concentrations. Two others took the opportunity to comment that the honeycomb grid in the RAM model is not consistent with the specification for receptor locations in Section 5. One of these commenters also requested that provision be made in the RAM model for user input anemometer height and user input wind speed profile.

There were seven comments concerning the urban versus rural typing scheme described in Section 5 of the guideline. In general these comments stated that the criterion had not been sufficiently validated or documented, that is, that the procedure had not been tested anywhere other than St. Louis. One of these comments requested guidance on what to do if there is a mixture of rural and urban conditions within the three kilometer radius. Once commenter stated that "urban" conditions may sometimes be appropriate in rugged rural terrain, while another comment stated that the guideline should acknowledge that the plumes from tall stacks are not subject to the same dispersion conditions that occur in the urban surface boundary layer, and, therefore, recommended that the Pasquill-Gifford coefficients be used until research can ascertain the correct dispersion curves for tall stacks in urban areas. Once commenter stated that he agreed with the EPA criteria for urban versus rural models.

There was a group of comments on the use of decay downwind for sulfur dioxide modeling. These commenters felt that sulfur dioxide decay should be included in all models intended for use after more than a few hours travel time. A somewhat related comment concerns the conversion to standard temperature and pressure that is provided for in the VALLEY model but not in other models. It was simply pointed out that the EPA model set was not consistent for this reason.

Models for nitrogen dioxide are technically screening models. Nevertheless, it is convenient to discuss them at this point. Four commenters suggested that an approximate first order chemical reaction for the transformation of nitric oxide to nitrogen dioxide should be provided for in the models. Three commenters pointed out that the data are seldom available to apply the ozone limiting method suggested by EPA. Regarding EPA's suggestion of the use of nitrogen dioxide to nitrogen oxide ratios, one commenter requested that EPA give examples of specific procedures. Another commenter suggested that the ozone limiting method should be allowed in rural areas as well as in urban areas if adequate data exist. Two participants suggested that the "algebraic" technique is more representative for rural nitrogen dioxide concentrations. One commenter suggested that the EKMA model should not be recommended until it has had the same degree of validation as EPA requires for other models. Another commenter simply said that point source screening models are not applicable to the nitrogen oxide problem.

One commenter addressed models for visibility impairment even though these models are not considered to be "recommended" models. He maintained that the weakest link in visibility modeling is the inaccuracy in the basic diffusion calculation.

Four commenters addressed the need for a model to obtain annual average concentrations for rural areas with significant terrain variations. Two of the four suggested that the ISC long term model be used for this application. Another said that a rural CDM version should be made available. Three other commenters pointed to a need for additional guidance in modeling fugitive particulates. One of these asked more explicitly whether ISC model can be used without gravitational settling or ground deposition.

There were half a dozen comments about the use of fluid modeling. Three of these stated that physical modeling would be the preferred method when adequate data are available. Another stated that the area of applicability of physical modeling is limited to tens of square kilometers not to a few square kilometers as stated in the guideline. The other two comments simply said that more guidance was needed on this subject in the guideline. On the related subject of downwash, two commenters said that the guideline should provide more of a discussion, including discussion of the so-called refined models. This commenter had in mind short term impact of multi-source situations, visibility, and long range transport. One commenter cited a need for models to be able to handle the merger of plumes 25-100 kilometers downwind of the respective sources. This commenter suggested linking upper winds with surface winds via a multi-layer ozone model as one approach.

Three commenters discussed alternatives to sequential models such as CRSTER and RAM. One of these suggestions was to use the model PTMAX or PTPLU to identify the most unfavorable meteorological conditions and then to determine separately the frequency of occurrence of these meteorological conditions. The other suggested, as a way of reducing the number of hours which models calculations are needed in the sequential models such as RAM or CRSTER, one could screen the meteorological data to find periods of greatest persistence of wind direction and then run the detailed models on the days identified by this persistence screening.

Lastly, with respect to the models actually listed as recommended models, there were four more comments. One said that the EPA should spell out the technical basis on which the recommended models were selected. Second, it was pointed out that, in a statistical comparison of the predictions of recommended models versus measured air quality, less than 20% of the cases agreed within a factor of two, and more than half the cases differed by a factor of five or more. Third, it was requested that the guidelines give specific literature references pertaining to validation of the CRSTER model. One person suggested that model changes should not be disseminated solely by in-house memos. Affected parties outside EPA could not be asssured of learning about all the model change memos. There was a suggestion that the screening model PTPLU be modified.

Screening and Refined Techniques for Complex Terrain

The second major topic on which EPA requested comments was the use of screening techniques and refined techniques for complex terrain. There were 57 comments in this category. First, there was a call for a more precise definition of complex terrain. One commenter said that more guidance was needed. Another commenter said that the term "complex terrain" requires definition, and suggested that complexity should be considered a function of elevation above stack base and stack top, plume height, and distance from source to receptor. A third person asked for clarification with regard to multi-source situations where stacks are at different heights. A fourth said that stacks shorter than a certain height should not be considered in determining flat versus complex terrain. One comment pointed out the need for a refined model for complex terrain, and requested EPA to commit itself to developing a complex terrain model as soon as possible. Another asked that EPA put in writing the criteria for acceptability of complex terrain models wherever derived. Another asked specifically for criteria to be spelled out for the use of the VALLEY model.

Most of the comments under this section concerned the VALLEY model. One commenter stated that the VALLEY model does not met EPA conditions for recommended models. That is, it is not "state-of-the-art." Another expressed his belief that EPA has required the use of VALLEY in a subjective manner. Once consultant stated that neither VALLEY nor COMPLEX I nor COMPLEX II should be used. It was pointed out that the "recommended" version of VALLEY differs from the version which was validated by Burt. This commenter stated that there recommended version gives concentrations several times higher than the validated version. There were several comments to the suitability of the VALLEY model as a recommended model. It should be pointed out here, perhaps, that the guideline refers to VALLEY as an acceptable screening model but not as a "recommended" model. However, the participants expressed a need for some kind of "recommended" model for regulatory decision making, and in fact, VALLEY usually is regarded in such a role. Several commenters cited the report of the National Commission on air quality and stated that continued use of the VALLEY model in his present form is inconsistent with the recommendations of that body.

Several commenters thought that the usual default values of inputs for VALLEY were too stringent, and therefore, should not be automatically required. One commenter stated specifically that the use of Class F stability is inappropriate in the rugged terrain of the Appalachian Mountains or the Ozarks. Another suggested the use of Class E stability is a reasonable worst case. A third stated that the use of VALLEY with Class F stability and 2.5 meters per second wind speed for six consecutive hours is hypothesizing input at its most conservative limits and should not be done. Another said that one should always check the validity of the inputs to VALLEY by reference to real data. And finally, one commenter said that the VALLEY model should assume the wind is always blowing away from the hills any time F stability is encountered.

A group of eight comments proposed various modifications to the VALLEY model. The first of these comments suggested that VALLEY should be modified to accept multiple sources and modified to accept a rectangular grid system. Other proposed changes were; first, to include bouyancy induced dispersion; second, to include only partial instead of full plume reflection from plant elevation; third, to make the wind speed profile consistent with other models; fourth, to modify the model so that mass is conserved in stable conditions; and, finally to incorporate wind shear, i.e. change of wind direction with height, or else the VALLEY model should not be used for regulatory purposes.

There were several further objections to VALLEY. Several commenters simply stated that they objected to the inclusion of VALLEY in the guideline, and the use of this model and setting emission limits. One commenter felt that the statement that the VALLEY model is acceptalbe precludes the use of more apporpriate techniques such as tracer studies. Another commenter stated that there is no technical basis for the choice of parameter values used in the VALLEY screening model. Another requested that the guideline clarify EPA policy on the use of the VALLEY model feature relating to correction of concentrations to standard temperature and pressure. This was also noted as an area of inconsistency between various EPA models, since VALLEY is the only model which provides for making this conversion. Another commenter suggested the use of a correction factor to convert 24-hour concentrations predicted by VALLEY to 3-hour maximum concentration for setting short term emission limits. This comment was intended to apply also to the models COMPLEX I and COMPLES II. Finally, it was suggested that further validation of the VALLEY model was required for the situation of narrow, forrested valleys.

There were eight or nine comments on EPA's proposed models COMPLEX I and COMPLEX II. One commenter said that these models were based on an inadequate technical foundation and that it was better to have no guideline model at all, i.e. no consistency, than to have an unvalidated model for the purpose. One commenter favored the use of sequential models but not necessarily this model. Again concerning specifically the models COMPLEX I and COMPLEX II, one commenter requested that EPA publish whatever validation studies it had on these models. The State of Utah commented that these models are too detailed to be listed as screening models. Two groups commented that the model COMPLEX II is too conservative, and specifically that it may give higher concentrations that VALLEY. The argument continued that, since VALLEY is known to be conservative, COMPLEX II is not appropriate. One commenter said that the terrain correction in COMPLES I and COMPLEX II needs a better exlanation, and then, upon hearing it explained by Dr. Turner, the commenter objected and said that it was inconsistent with the treatment in CRSTER. Finally, one commenter supported the use of COMPLEX I at least until such time as EPA's small hill investigation produces a better model.

There were 15 general comments on complex terrain. One commented that procedures other than the screening models should be found to handle PSD and air quality maintenance planning. Another said, pending study of new models, EPA should allow the use of any "state-of-the-art" model. Another comment similar to an earlier comment was that EPA, perhaps in cooperation with the American Meteorological Society, should provide refined models for complex terrain. One commenter suggested the use of flat terrain models until reasonable error bands on complex terrain models can be developed. Two commenters stated that refined models are, in fact, available. They cite several models that have been shown to be better than VALLEY. One person said that EPA should make use of complex terrain studies conducted by NOOA, and that these studies should be used to improve concentration predictions obtained by the EPA screening models. Two commenters suggested that the EPA field experiments should be refocused on two dimensional rather than isolated three dimensional features. Six of the 15 general comments on complex terrain suggested alternative approaches. One suggested that meteorological input from more than one station should be allowed. The use of National Weather Service upper air charts was also suggested by the same commenter. One person recommended the use of physical modeling for complex terrain. Two people suggested using CRSTER; one suggestion was for flagging the output for any receptors above stack height, the second suggestion was for CRSTER to be used with a modification to include terrain lifting. One specific terrain lifting algorithm recommended for use with COMPLEX I and II was to use plume displacement equal to 35% of the effective plume height when the terrain height was near the effective stack height. This commenter recommended that the on-going studies by EPA be used to establish the displacement in the vicinity of an isolated hill. Finally, when the terrain height was between stack height and the effective plume height the recommendation was use COMPLEX I or COMPLEX II.

There were also a number of other comments on the use of screening models, not necessarily restricted to screening models for complex terrain. One of these comments was to the effect that the use of screening models is endorsed providing the results from screening models are not the sole basis of regulatory decisions. Another said that screening models are conservative enough already, and that choosing extreme inputs compounds this conservatism. One state agency suggested a need for a formal definition of "conservative" in this context. Suggested definitions were: "conservative estimates are estimates which equal or exceed the highest that may be expected", or "conservative estimates are estimates that exceed the highest concentration that would be measured had a monitor been placed at that location." Another comment went farther to say that the use of the highest short term estimate obtained from the screening model as a basis for emission limits is indefensible. A third said that the use of screening models for increment tracking may give the erroneous impression that too much PSD increment has been used. This particular state agency would prefer a refined model whenever possible. One comment requested EPA to also define a level of impact for which refined modeling will not be required. Two said, in effect, that the EPA screening procedures should be updated. One of them, being more specific, suggested that PTPLU should be included in the Volume 10 procedures. Several said that EPA should give more guidance on screening models for such topics as inversion break up fumigation and lake shore fumigation, long range transport, fugitive dust, visibility, and particulate emissions from cooling towers.

Requirements for Meteorological Data

The third topic on which EPA requested comments was requirements for the modeling data base: emissions, meteorological, and air quality. The comments concerning meteorological inputs will be discussed here. Comments on other aspects of the modeling data base will be discussed later in the report.

Length of Record

The 1978 guideline states a preference for multiple years of meteorological data. The 1980 proposed guideline states that five years of data are required if National Weather Services airport data are used and that consecutive years from the most recently available five-year period are preferred. More than 20 commenters objected to either the need for a five year period of airport data or EPA's preference for the "most recent" data. The five year requirement was felt to be too costly, especially for a small project. Many stated that EPA had not provided adequate justification for requiring a five year meteorological data base. Most felt that one year was adequate, although one commenter suggested a compormise at three years. Many of those who felt a one-year data period was adequate were referring to on-site data. Several commenters suggested a dual screening process, wherein a sequential model with one year of hour-by-hour meteorology would be used as a rather refined screening model. Only if concentrations calculated in this one year exceeded some threshold, say 80% of the limiting concentration, would it be necessary to go on to the full five-year modeling period. A few commenters challenged EPA's requirement for the "most recent" data. The difficulty seems to be that the National Weather Records Center has not routinely keypunched hourly data since 1964; therefore, anyone wishing to use "most recent" data must arrange for keypunching it. The commenters argued that EPA had not sufficiently justified this requirement.

There were two miscellaneous comments with regard to meteorological data availability. One pointed out that the Air Force Air Weather Service data would be suitable. The second asked EPA to set up a regional or nation-wide clearing house to facilitate identifying sets of data (suitable for modeling) that had been gathered in the vicinity of proposed new sources.

Temporal Representativeness

The 1978 guideline indicated that it might be necessary to limit modeling and subsequent analysis to a single year of data, if only one year were available. The guideline also suggested that using a single year of data might be justified if the temporal representativeness could be demonstrated by reference to a longer record from a nearby National Weather Service site. No further guidance was given. The 1980 proposed revision gives a particular procedure to be followed in demonstrating such representativeness. Twenty-seven comments addressed the proposed procedure. Approximately 15 of these said the test was unworkable. One commenter felt that the representativeness of on-site meteorological data should be determined on a case by case basis. A nearly equivalent statement was that the decisions regarding representativeness of data should be made by regulatory agencies having the new source review authority, not necessarily EPA. Two participants suggested that the guideline should clarify that it is the National Weather Service data from the test year that is to be compared with the averages for the other four years. Two of these suggested that the observations from the test year should be included in the five year average against which the test year is compared. One or two participants cited examples illustrating why the meteorological data from a site should be different from the meteorological data recorded at an airport. However, it seems to be that these people were mistaking the nature of the proposed test.

Two participants stated that in a question of on-site versus off-site meteorological data, the on-site data should always be preferred. One of the two stated that National Weather Service measurements, presumably referring to airport measurements, are not suitable for air quality prediction. Two others said that in complex terrain, one year of on-site data is always preferable to longer periods of off-site data regardless of how many years record might be available off-site. A third participant endorsed the use of on-site data; however, he noted that there would be a considerable difference in CRSTER results using on-site versus off-site data. In contrast, two commenters suggested that on-site meteorological data are not needed in most situations. With regard to determining the representativeness of on-site data, one commenter asked for guidance on how to determine representativeness when the only nearby data are the data gathered on site. Another similarly suggested a need for EPA to recommend specific alternative courses of action for data sparse areas. One person suggested that the problem of non-representative years could be solved by using a short period of on-site data, to correct the off-site data in such a way that the longer period of record of off-site data could be used to synthesize an equivalent period of on-site data. Another person cited a need for EPA to make a statement regarding the acceptability of a second, third or fourth year of data from on-site. The representative of the Nuclear Regulatory Commission stated that NWS data are always of uncertain representativeness and suggested that the guideline should present criteria for determining representativeness of NWS data.

Application

One comment suggested that, when worst case meteorological conditions must be evaluated, consideration should be given to probable combined worst case conditions. Another comment suggested using five one-year STAR data sets, not a single 5-year average data set to evaluate annual impact. A third comment requested guidance on use of the direction randomization feature of the EPA meteorological pre-processors when on-site winds are already reported to the nearest degree or tenth of a degree of azimuth.

Comments on Appendix C

There were approximately 60 comments on various aspects of Appendix C. Six people endorsed the measurement of sigma theta on site and the use of this parameter to determine sigma y directly through Pasquill's formula, and the use of this sigma y in a split sigma approach in place of the standard sigma y of the determined Pasquill class. One of these posed the question "Are there any circumstances when a regulatory agency should insist on sigma theta measurements despite the expense to an applicant?" Three other comments addressed the averaging time associated with the determination of sigma theta. Two of these simply requested EPA provide guidance on the averaging time associated with the sigma theta, and a third made specific suggestions for the use of 15 minute sigma thetas. Another nine comments addressed the use of z_o to adjust the values in Tables C-1 and C-II. Six of the participants simply requested that the guidelines provide some kind of guidance or table of z_o values for use in modifying Table C-I and C-II. Two participants gave suggestions for a table z_o to be used with these tables. One person pointed out that the guideline is confusing as to whether the measured or the tabulated sigma value should be adjusted. This person also questioned whether any z_o adjustment was appropriate at all because the larger wind angle fluctuations associated with the larger z_o values properly reflect the increased dispersion rates which would be found at a rougher site.

There were another seven comments pertaining to the use of sigma theta as a means of determining sigma z. One person expressed the hope that sigma theta will not be required as an estimator of sigma z if sigma phi has been measured. The Nuclear Regulatory Commission suggested that sigma theta should be used only if the wind speed is greater than 1.5 meters per second. Another participant stated emphatically that sigma theta should not be used to define sigma z but sigma phi or delta T is preferred instead. Another supported the use of delta T for determination of sigma z and two suggested that the measurement of net radiation on site should be an option. One commenter suggested an alternative procedure for determining stability classes at remote sites.

Another group of eight comments expressed dissatisfaction with using any type of on-site wind fluctuation data to determine stability categories. The prevailing feeling in this subgroup was that EPA had not provided sufficient validation for the wind fluctuation approach to stability class determination. One participant suggested that EPA should develop a way of estimating these parameters from airport data. Another pointed out that the use of on-site stability would be inconsistent with nearby airport data. And one person said that the recommended sigma phi equivalences in Appendix C-III are incorrect. Another person pointed out that the recommended procedures, including the adjustment in Table C-III, will result in large jumps in stability classes when the sun rises and sets. Still another comment stated that the stability determination scheme given in Table C-III was intended by its authors to be used only in determining sigma z in a split-sigma approach. One person suggested that sigma theta and sigma phi over rough terrain should be measured at a height only five meters above the average of trees or roughness elements.

A group of eight comments was related to on-site tower measurements. Two of these stated that if tower winds were availalbe, then the wind speed at stack height rather than the wind speed near the ground should be used in determining Pasquill stability categories. One participant stated that towers are very costly and suggested that EPA develop other ways to estimate the winds at heights of 200 meters or more. Several participants recommended measuring wind speed at at least two heights on a mast. Two people went further and suggested that when wind measurements are available from two different heights on-site that the wind speed profiles should be developed directly from the wind shear data measured at the site. One person suggested that it is not appropriate to require towers taller than 40 feet and that the data from two levels be used to estimate winds at higher levels. He indicated there is not enough benefit to justify the cost of erecting a taller tower. One state agency requested guidance on two topics, one of them being the use of vector averages, the other being use of wind data if it is collected at more than one level.

There were half a dozen suggestions for modifications to the instrument specifications given in Appendix C. Four of these concerned the specifications for a vertical propeller anemometer for measuring vertical velocity fluctuations. One suggested that the EPA instrument specifications would result in a gross underestimate of the vertical velocity variance. The recommended specification for the vertical starting speed was 0.1 meter per second. One commenter suggested that the sampling frequency for a vertical propeller anemometer need not be as fast as every 10 seconds. He suggested that, if the propeller were mounted at 100 meters above ground, a lower sampling rate would be adequate. There was one request for guidance on quality assurance for on-site meteorological programs. And one comment that "hourly" averages should be based on not less than 15 consecutive minutes.

Finally there was a group of comments related to the determination of mixing height. Three of these comments simply pointed out that the mixing height is an item of information required for modeling, and felt that it should be added to the list of parameters that EPA requires to be measured on-site. Two people suggested using acoustic sounders for on-site determination of night time mixing depth, and two suggested the use of the Benkley-Schulman technique for interpolating mixing heights from hour to hour between radiosondes and also the use of this technique for cold advection cases.

Treatment of Calms

The meteorological pre-processor for the EPA "recommended" models that use sequential hourly meteorological data adopts a procedure for calm hours which assigns to these hours the unrandomized wind direction from the last hour of valid direction and assigns a speed of one meter per second. Unless the sky is heavily overcast, this procedure results in string of Class A hours in the daytime or Class F or G hours at night, both with only slight changes of wind direction. Class A normally results in highest ground level concentrations for tall stacks, while class F or G result in maximum ground level concentrations for surface sources. The lack of significant wind fluctuations in such hours is felt to be unrealistic because, in many cases, there is a great deal of meander associated with mean wind speeds as low as one meter per second. In the October 1980 proposed revisions of the modeling guideline, EPA proposed to remedy this lack of realism by casting out of any averaging period all the predictions for hours with calm winds if those hours made non-zero contributions to the longer average. Thirty-two comments were received on this proposal. One person said that the basis of EPA's proposal was unclear, while still another person requested further clarification of the procedure. One commenter stated that the proposed procedure should apply to calm hours in on-site data as well as to hours reported as clam by the NWS.

Two comments stated that the new procedure should be applied only to 24-hour averages, not to the three-hour averaging time. Another felt that EPA guidance should also state a minimum acceptable number of non-calm hours, and that if the number of non-calm hours was insufficinet, then the calculation for the entire averaging time should be considered invalid. One person suggested that calm hours should be excluded from annual average calculations as well as the short term calculations. Two or three commenters proposed 18 hours as the minimum number of acceptable hours for a valid 24-hour average. Two other participants suggested eliminating any claculated concentration for a calm hour if the calculated concentration for the calm hour exceeded the average of all the non-calm hours for that averaging period. Two people suggested that applying the EPA procedure to three-hour averaging periods may result in higher concentrations than if certain non-zero calm hours were not eliminated. These participants felt that all three-hour averaging periods having less than three hours of non-calm calculations should be discarded. Three commenters suggested that EPA should incorporate the procedure for casting out calm hours as a feature to be performed automatically in the recommended models.

A group of ten commenters was received which suggested that EPA, instead, should use a "double randomization" procedure for calm hours. The double randomization procedure presumably means that a random direction between 10° and 360° would be assigned to the missing hour and then the hour would be subjected to the plus or minus five degrees randomization which is applied to all normal hours. Two of these comments suggested a plus or minus 30 degrees range for the first randomization and two others suggested plus or minus 22-1/2 degrees. The other comments did not suggest a specific range for the extra randomization. One of the ten suggested that this double randomization procedure should be applied at most to one third of the hours in any averaging period. (That would not be more than one hour for the three-hour period, and not more than eight hours for the 24-hour period.) If a greater number of hours were calm, the commenter felt the whole period should be discarded. Still another commenter suggested that no prediction should be made for periods when less than 75% of the data are available. Finally, a group of four comments stated, in essence, that calms should not be ignored or discarded, particularly, as one commenter pointed out, in the case of input for segmented puff type models. Another participant said simply that calms should not be ignored, and that a better way should be developed. Another commenter who stated that a better method was needed said that the proposed method minimized unrealistically conversative predictions, but did not resolve the problem of modeling calms properly. Finally, one comment suggested human intervention to put in bogus data to replace the hours reported as calm.

Criteria For Acceptance Of Non-Guideline Models

Forty-seven comments were received on this topic, including comments on the related topic of tracer studies. Section 6.1 of the proposed guideline gives definitions of what constitutes a non-guideline model and outlines three different categories of review. Eight participants commented on the guideline definitions and the proposed categories. One person said that the definitions in Section 6.1 needed improvement. Another said that criterion b. is inappropriate, citing as an example a simple change of output format in a model which would otherwise produce the same numbers as a guideline model. Three others said that it was difficult to ascertain exactly which models were the "recommended" models that would require no demonstration. One of these people said it was almost impossible to find out which versions of the UNAMAP models are the acceptable versions. Another said, except for designating CRSTER as a benchmark for large point sources in rural areas, the guidelines is unclear about which model options are "recommended". Another specifically asked, in the case of wind data at two levels being used to determine a local wind profile as recommended on page 39 of the guideline, whether the use of this profile in a model means that the model is non-guideline and therefore subject to all the demonstration requirements of Section 6. Two comments stated that review category 1 should be broadened. The category, as proposed by EPA, says only output format changes or similar changes The commenters felt it appropriate to broaden this category to include changes in computer coding that change only the computation efficiency but don't change the results. Another commenter stated that he could see no difference between models in the review Category 2 and models in review Category 3, and there was a need for definition of major changes and minor changes. A second comment also indicated that EPA should clarify the difference between Category 2 and Category 3 review.

Finally one commenter suggested that the guideline should permit a decision that a "recommended" model is not appropriate based on computational expense. Six comments addressed quantitative limits for Category 1 demonstrations. One said that an agreement within 2% was unrealistic. Another that accuracy of models is to be preferred instead of agreement within 2% with an EPA model, especially in light of the fact that EPA models have no or poor verification. Another said, equivalently, that the proposed 2% agreement test is irrational because EPA models are of unknown validity. One said that no demonstration of equivalence at all should be required for non-guideline models with such elementary changes as would put them in EPA's proposed Category 1.

Five oral comments and one written comment expressed the thought that EPA's demonstration criteria posed a double standard. One of these comments said that EPA's procedure serves the objective of consistency but does little to improve the estimates of true impact. A companion comment stated that EPA, in effect, can put forth an unvalidated model, but applicants have to gather extensive data to prove their alternative model is better, and that there is little evidence EPA has done anything to improve the accuracy of models. Another person said that EPA has provided no data showing that the recommended models are best, and that the difficulty in meeting EPA demonstration standards will stifle the development of new models. Finally, one said that non-guideline models should not have to show computational equivalence; that models may be compared to guideline models but not validated by such a comparison. Two of the commenters felt that, until objective model performance measures are developed, no demonstration at all should be required. Two other comments suggested as a criterion a much larger band such as 10% or 20% or 2 ug/m^3, not 2% agreement. Another person said that non-guideline models should be permitted as long as they are equal in performance to the guideline models. Another said that the guideline models should have no special status since they are unproven. Still another questioned whether the EPA models themselves could satisfy the criteria set forth later in Section 6 for acceptability. One participant indicated that EPA should be required to justify their present models. Another said that such demonstration requirements should be required for all new EPA models prior to inclusion in the guideline.

Another 16 comments addressed demonstration studies, including tracer studies. One comment, which is perhaps more appropriate to the preceding section, stated that EPA review of modeling should be based on the modeler's statement of whether the guideline techniques have been used. Two people stated that the criteria for demonstration studies are so strict as to be virtually impossible to satisfy, and that EPA seemed to be erecting barriers to the use of non-guideline models. As an example of the difficulty of demonstration studies, yet another commenter cited the requirement that an appropriate model be used to position monitors, but by definition no prior appropriate model exists. Another comment suggested that several terms in Section 6.2 need explicit definitions. These are "appropriateness," "similar," representative" and "correlation". The proposed guideline suggests in-stack monitors would be necessary in order that the accuracy of emissions be comparable to the accuracy of monitoring. One comment stated that an in-stack monitor would not be needed in the case of an oil-fired powerplant. One of the commenters suggested a chemical mass balance as a means of making particulate data more useful.

One participant suggested that the portion of Section 6.2 dealing with tracer studies needs to be clarified. The same set of comments indicated that tracer studies should be designed to reflect worst case conditions and designed to measure the location of maximum concentrations. A different commenter felt that tracer studies should be listed as one type of physical modeling under Section 3.2. Finally, one commenter warned that users should be made aware of the expense of tracer studies and the possibility that the results may not be acceptable nor provide the final answer.

Section 6.2.3 of the guideline addresses the circumstances under which a diffusion experiment conducted at one site can be used as validation of a model proposed for a different site. Three of the four comments in this area felt that the criteria for the use of off-site data were too restrictive and/or too vague. It was suggested that the guideline should list these criteria as considerations to be addressed, not as hard and fast rules. Finally, one participant suggested two other factors that should be considered in comparing models for valley situations. These are the direction of orientation of the valley and the thickness of the drainage flow.

Placement of Monitors to Support a Model Evaluation

EPA requested comments on the appropriate placement of monitors to support a model evaluation. Three participants responded. One suggestion was for a single monitor corresponding to the maximum concentration expected for each averaging time plus one mobile monitoring system. Another stated that it can be as few as four locations and that the choice of locations must begin with an agreed measure of performance. This comment also added that EPA should publish evaluation studies for the EPA models. One comment suggested a need for at least four or five samplers for each distance of interest.

There were another eight comments, which, while not specifically addressing the placement of monitors, are more related to model demonstration field experiments than to any other topic. One said that all the data gathered prior to EPA's announcement of PSD monitoring guidelines should be allowed even if the data do not strictly meet PSD criteria. Another stated that the quality assurance procedures for demonstration studies might need to be more rigorous than the procedures contained in the PSD guideline. One comment stated that the meteorological data for tracer studies should be adequate and representative of the transport and diffusion processes affecting the plume. Three others suggested the use of separage validation for plume rise sub-models. Different instrumentation such as COSPEC, LIDAR and/or airborne tracer studies could be used as a means of validating the plume height sub-component of the model.

Performance Measures and Standards for Models

There were 13 general comments and several specific comments on standards or measures of performance for models. The Utility Air Regulatory Group endorsed the use of model performance criteria, stating that the current guideline has a poor technical foundation because performance standards do not exist, and urged that EPA give top priority to developing objective performance measures. The El Paso National Gas Company suggested that any comparison of models should be based on performance criteria, and that EPA should place high priority on the development of such criteria. The Salt River Project stated that before criteria can be applied to non-guideline models, the guideline models must first be evaluated in quantitative terms, and then the performance measures must be applied to all models, both guideline and non-guideline. The American Petroleum Institute stated that EPA has an obligation to propose performance standards. They further suggested that one way of proposing such standards is to document the performance of guideline models as bench marks for acceptable model performance and that this should be done across the range of conditions for which the models are recommended in order that users can understand the uncertainties.

The Virginia Electric Power Company suggested that a rational performance evaluation procedure should be developed, and that model developers or EPA should document the sensitivities of their models to variation of inputs and document the model error limits. Finally, one other comment suggested that statisticians should be represented in the development of performance measures for model evaluations.

Five different comments suggested possible performance measures. One emphasized that, not only is good case-by-case agreement needed, but also that the highest predicted concentration should occur under the smae meteorological conditions as the highest observed concentration. A second comment suggested the following:

1) Agreement of time independent peak concentration within a factor of two of the instrument error.

2) The absolute difference period by period between predicted and observed should not be large.

3) The ratio of predicted to observed should always be in the range of 0.8 to 1.5.

A third comment simply suggested that EPA look at the work being done by the Electric Power Research Institute. Another suggested the following statistics: (1) The error relative to observed concentration, (2) percent over-predictions and under-predictions, (3) percent within a factor of 2, (4) then the variance and shape of the statistical distribution of (observed minus predicted) over observed, and (5) behavior of the statistical distributions at the upper concentration end. One participant suggested a particular mathematical form for representing the high end of the spectrum of concentration versus frequency.

Three comments echoed sentiments already reported under the topic of Demonstration Studies. One of these is that CRSTER should not be the benchmark. Another stated that performance standards should not be required of any model because they haven't been required of EPA models. A third said that the guidelines should contain references describing the detailed validation results for the EPA models including some event by event comparisons.

Model Evaluation and Calibration

There were twenty-two comments related to Section 8 of the proposed guideline. Two participants strongly encouraged comparing model results against actual air quality data, and stated that calibration should be used to improve model accuracy. A third stated that model evaluations should be made in all situations where post-operational monitoring is required. This person also stated that summarized evaluation results would be a useful addendum to the guideline. One commenter, however, was opposed to model calibration, especially when many emission sources are unquantified. And one commenter simply said that a thorough discussion of model accuracy was needed in the guideline. Two commenters felt that validation of the guideline models should be accomplished. One of them noted that the state of validation of EPA models is not much better in 1980 than it was in 1977.

Two comments specifically addressed statements in the proposed guideline. One proposed that the statement "there is little justification for using a model that accounts for less than 50% of the variation of the measured concentration" should be changed to say "there is little justification for using a model that accounts for less than 50% of the significant concentrations". Furthermore, a definition of "significant" should be provided. The other said that model applicability should be judged on the basis of prior comparison with actual monitored data.

Two comments raised the issue of sensitivity analyses, one recommending that model sensitivities should be discussed in the guideline, in particular sensitivities to the uncertainties in transport wind and uncertainties to parameters controlling the sigmas. The other commenter said that EPA should perform evaluation and sensitivity analyses for the "recommended" models. Another said the uncertainties in the guideline models should be given in an appendix to the guidelines. In support of a statement that the uncertainty of models should be emphasized, one commenter cited that uncertainty in plume rise and the scatter of true sigma values about the predictions of the published curves are both sources of model error which should be emphasized. One commenter objected to focusing on errors in the input data when the weaknesses of other model assumptions are ignored. This same commenter took exception to the guideline statement that establishing model sensitivities is a responsibility of the model user. He felt that establishing precision, accuracy, and sensitivity was a responsibility of the model developer even if the developer of the model was EPA.

Four participants made comments on some mechanical aspects of model validation. Two of these emphasized that all the data inputs should be averaged over the same averaging time. Such a requirement would obviously eliminate the use of National Weather Service data as a source of meteorological input for validations. Another took exception to EPA's call in the guideline for evaluating the entire model, and said it should be possible to validate the components of the model separately. The last of these suggested that the ad hoc adjustment of sigmas and plume rise is most effective in the case of short-term concentrations.

Finally, there were four comments on the related topic of confidence intervals. All four of the comments endorsed the concept of presenting confidence intervals along with model results. One stated that computer output should not simply be compared with standards. Another stated that the public is best able to judge risks when modeling is presented as best estimates with associated confidence limits. A third simply endorsed error bands. The fourth said that the guideline is unsatisfactory because error bands are not presented. He went on to say that if models do not perform acceptably, they should not be used as a basis for setting emission limits.

General Comments on the Guideline

There were a large number of comments which do not fit well into any of the other categories. Two commenters felt that the 1980 proposed revisions of the guideline represented an improvement in several areas. On the other hand, one commenter stated that the guidelines were so deficient in general approach that the whole guideline should be revised and reproposed. Another participant stated that the guideline should include diagrams. One person objected to the availability of model code in FORTRAN as a criterion for model availability. This person said the code

BASIC is available on most big computers and is more understandable to noncomputer people. One comment stated that the guideline should perform a critical review function (similar to the criteria documents). The comment went on to say that the guideline should review the available modeling techniques before making any recommendations. The Department of Housing and Urban Development suggested that it is indoor pollutant concentrations which should be tested against air quality standards, and suggested that EPA account for defferences between indoor and outdoor air quality in testing model output against air quality standards.

The guideline states a special requirement for sources located within ten kilometers of a class I PSD area. One commenter suggested that this requirement is equivalent to the establishment of the buffer zone outlawed by the Clean Air Act Amendments. In discussing nonguideline situations, the 1980 proposed guideline recommends, on page 10, early discussions among regional office staff, appropriate Federal Land Managers, the state and local agencies and industry representatives. Two comments stated that the Federal Land Managers should have no role in such discussions.

There were five comments on the proper role of models in the decision-making process. One stated that the use of models as a primary tool in making decisions is not inevitable. He recommended using modeling as a guide in setting generic emission limitations but saw no need for each of several thousand applicants to have to apply modeling individually. The comments of the General Motors Corporation objected in principle to the use of modeling in regulatory decisions. GM believes in relying on generic source specific emission limits. Another stated that model results should not be the sole tool for making judgments. The fourth recommended that EPA adopt a philosophy of using models as tools in decision-making. Finally, one said that models are not perfect, therefore one must be cautious in their use.

Four comments raised an issue which had been raised in 1977. These called for establishing some kind of technical review panel to resolve disputes over technical issues in modeling. On a related issue, three comments urged that the EPA models be subjected to some form of peer review, before the models are used in a regulatory setting. Three comments called for written documentation whenever case by case decisions have been made by the staff at Research Triangle Park or by regional administrators. One also said that past decisions should be "grandfathered".

Consistency

There were sixteen comments on the apparent conflict between modeling consistency and modeling accuracy. The EPA viewpoint on this topic was expressed succinctly in the Washington, D.C. meeting by Richard Rhoads in his welcoming remarks. He spoke of the dual goals of accuracy and consistency and indicated that the EPA valued consistency the more highly of the two.

The opposing viewpoint was stated most dramatically by the representative of the Minnesota Pollution Control Agency in the Chicago meetings when he said, "It's the real world we are concerned with, not consistencies. If EPA can't see that, I think we are in trouble." ONe commenter asked the rhetorical question, "Since models are inaccurate, what is the point of trying to be so consistent in precision?" Five other commenters made such points as: (1) realism is the key; (2) consistency should not be valued so highly that it produces inflexibility; (3) consistency should not be an overriding goal; (4) accuracy should come first; and (5) flexibility is recommended, especially together with the use of on-site meteorology.

Four sets of comments stated that EPA had misinterpreted the public's desire for consistency; that is, the EPA viewpoint is not what the agencies want. Another comment was that the EPA concern for consistency should not be followed to the point of accepting consistently inferior results. One participant said that the consistency should mean uniform procedural steps such as (1) requiring Regional Administrators to put their decisions in writing (2) providing a summary of materials used in making decisions, and (3) making a record that is reviewable as in the Administrative Procedures Act.

Three commenters expressed the opinion that the EPA emphasis on the use of "recommended" models is undesirable from a scientific point of view; that the EPA position on consistency would produce a series of identical "clone" models; and that the EPA view of consistency would prevent professional meteorologists from exploiting new information. One person stated that the desirable model characteristics are: (1) cost effectiveness, (2) scientific credibility and, (3) reasonable accuracy. Another comment suggested there were three essential elements to consistency: (1) the use of representative meteorology, (2) consistent criteria for urban vs rural characterization and (3) choosing models from an approved list. Another participant stated he would prefer case-by-case consideration of the physical assumptions. One comment said that the guideline rigidity negates the value of having the modeling effort directed by "competent individuals with broad range of experience in air pollution meteorology." Still another comment supported a statement in the guideline itself: "In all cases the preferred model is that which best simulates atmospheric transport and dispersion in the area of interest."

One set of comments indicated that the inconsistency was found at different levels within the same EPA region. For example, the state of Maine may do things differently from the way EPA Region One would like. This commenter indicated that consistency within a single EPA region is highly desirable; on the other hand, consistency from EPA region to EPA region is not necessary and might even be inappropriate. Another set of comments indicated consistency can be achieved adequately without the rigidity of the present guideline. This comment also said that consistency should mean a uniform approach in determining which model is best. Another commenter stated that the basic problem of consistency lies in the model use and interpretation, and still another commenter mentioned that the guideline should provide consistency in the use and interpretation of model results. A participants expressed concern that the guideline tends to "cast in stone" techniques that are not economical and not technically appropriate. Still another commenter felt that the current EPA policies seem to discourage model development. One lone commenter stated that the guidelines did not go far enough in the direction of achieving consistency.

One commenter suggested one method of diagnosing regional consistency is to have a "modeling round robin." Two other felt that consistency would be enhanced by EPA having a system to keep all users current on changes to UNAMAP programs and that special technical conferences should be held to transmit modeling information to the state agency level.

There were a number of comments on the way EPA revises its models and on procedures for letting the world outside of EPA know about the changes. One state agency requested an opportunity for better communication and closer collaboration between state agencies and EPA. One EPA regional meteorologist stated that an air quality modeling clearing house should be established. A comment pointed out the long time lag between EPA's promulgation of new models and the time that a state is able to use them. An architectural engineering concern representative suggested that EPA publish a model newsletter to overcome the difficulty people outside EPA have in keeping up with EPA thinking. Two state and local agencies stated that if the EPA's method of revising the guideline is going to be through workshops, then state and local agencies should participate. One participant indicated that waiting for the National Technical Information

Service to make models available is too slow. He also suggested a clearing house was needed to rapidly disseminate model revisions and corrections. Another suggested that EPA should forward memos on modeling to everyone who is a user of the UNAMAP system as soon as possible so that all users would be on the same footing. One participant recognized possible inconsistnecies, saying that the guideline models should be clearly identified, and that he would prefer that the models be listed in tabular form. Another said that the 1978 version of the guideline should prevail until after public review of the 1980 proposed changes had been completed.

Comments On Section 2

There were two general comments on the interpretation of modeling results for determining compliance with air quality standards. One industrial commenter stated that there was a disturbing trend toward placing too much reliance on numbers. He suggested EPA should establish a policy such that meaningful decisions can be made without paying overly simplistic reverence to computer generated print-outs. One representative of a state agency suggested that no more than two significant figures should be used. Also, one of the EPA regional meteorologists requested that the definition of "ambient air" appear in the guideline.

The representative of one state agency said that the guideline needs to clarify the meaning of the highest-second highest concentration when multiple years of modeling are used. Another comment on the use of multiple years of meteorological input in modeling suggested that the median estimate not the highest estimate of any year of the multiple years be used as the basis for assessing compliance with short term standards. He was suggesting, in effect, that compliance with short-term standards should be met only in typical years, not in every year. A third commenter said that the policy implication of requiring modeling a five-year meteorological period is the same as the policy implication of requiring a tighter receptor grid. He called for assessing compliance with standards on a statistical, or risk, basis.

There were seven other comments addressing EPA's policy of setting short-term standards such that the limiting concentration may be exceeded only once per year. One comment simply stated that this was a questionable concept. The other six suggested that the level not to be exceeded more than once per year should be replaced by a level having a certain probability of exceedence. For instance, the standard might be stated such that the 95th or 97th percentile of the short-term frequency distribution could not exceed the stated value. Most supported either the 95th percentile or the 97th or 98th percentile. One suggested the 95th percentile for twenty-four hour standards and the 85th percentile for three-hour standards.

There were a half a dozen comments on Section 2.2.2 of the guideline. One commenter simply requested that the term "significant impact" as used in the first paragraph of section 2.2 and again in section 2.2.2 be defined. Another person requested guidance on how to interpret the consumption of PSD increment from multiple sources. He asked whether the maximum short-term increments from different sources should be added together regardless of their location. Another suggested specifically that the guideline should read that the short-term PSD increment is obtained by subtracting the highest-second highest receptor concentration in the baseline on a receptor-by-receptor basis from the projected future second highest concentration. It should be noted that this is not the same as looking for the second largest difference at a given receptor between the present concentration and the future concentration. Two commenters opposed the EPA requirement for modeling sources located within ten kilometers of a Class I area even if the emissions are normally considered to be not significant. One of them suggested, however, that, if such modeling is required, then the criterion for impacting a Class I area should be changed to read "one microgram per cubic meter or one half of the available increment, whichever is larger."

General Comments On Section 3

Most of the comments pertaining to Section 3 of the proposed guideline were discussed under the heading of Recommended Changes to EPA Guideline Models. However, there were a few other general comments. One stated that models should be unbiased. Another asserted that the models recommended in the guideline are biased toward higher predictions, and that the new procedures increase the bias. A third suggested that the main criterion for the suitability of models should be accuracy. Four other commenters suggested that the cost of running various models should be considered when recommending model suitability.

One participant made a very specialized suggestion regarding the choice of single-source versus multi-source models for particulates in sulfur dioxide. The suggestion was that the model CRSTER be used for a single plant when the separation between stacks is shorter than the minimum downwind distance for a maximum concentration using the model PTPLU. For stacks separated by a greater distance he would recommend the model MPTER.

Two other comments addressed special modeling situations. One requested guidance on two types of special situations. One was downwash modeling, the other was modeling for fugitive particulates. Regarding the latter, the commenter said the guideline should discuss the inputs to fugitive particulate models, such as what types of sources should be considered point sources, areas sources, or line sources.

Requirements For The Emmissions Data Base

Section 4 of the proposed guideline addresses input data requirements. The requirements for meteorological data have already been discussed in an earlier section. Comments made concerning the requirements for source data are summarized in this section. There were four general comments concerning emission rates for single-source modeling. One of them said the guideline was ambiguous on pages 27 to 28. Another requested, in case there is any difference between actual maximum emission rates and allowable emission rates, that the actual rates be used. Another commenter questioned whether there is any benefit to be found in searching for the precise load level that produces the maximum impact. This commenter felt that modeling at the maximum rate or the allowable rate is sufficient to find the maximum short-term concentration for most applications. Finally, one commenter requested guidance on emission rates for SIP analyses, and one commenter said that modeling for more than one load level is burdensome.

Six comments suggested a statistical approach to choosing the maximum operating rate for modeling. One of the comments said that worst case load and meteorology should not automatically be used together in setting emission limits. Another comment said that one should not model the worst load as a constant. This comment suggested adapting statistical procedures. A third said, from industry's point of view, the most serious problem in modeling is a failure to use the joint probability of occurence of unfavorable load and unfavorable meteorology. Another commenter suggested modeling the 95th percentile load, that is the load which is not exceeded 95 percent of the time. Three comments specifically advocated the use of EPA's "Ex Ex" technique, and one simply stated that a statistical approach to expected exceedences should be used by EPA in the guideline.

There were fourteen comments concerning load levels for multisource modeling. One person suggested using screening models to find the worst load individually for each of the multiple sources and recommended that these worst-case load levels be used in all the multi-source modeling. Other possibilities were the use of one hundred percent load for all sources and use of the most likely load level for all sources. Two commenters suggested use of the probable maximum emission rates in multi-source modeling. This commenter suggested that the use of the maximum probable emission rates for all sources in multi-source modeling will lead to over-predictions. Two commenters requested that EPA give a more complete definition for the phrase "statistical maximum operating rate." Another comment stated that multiple loads should not be required in refined modeling of multiple sources if screening models show lower loads to be no problem. Another person simply said that, for use in multi-source modeling, one should not use all the sources simultaneously at a hundred percent load. Another comment complained that the guidance is not clear about what loads should be used for multi-source modeling.

The proposed guideline suggests using annual average emissions for area sources, but one commenter rejected this approach if probable maximum emission rates are to be assigned to point sources. Two commenters suggested, in the case of PSD modeling, that actual emission rates for the PSD "baseline" sources should be used. Another person thought that the requirement for multi-source modeling should either be omitted from the guideline or else a mechanism must be provided by which an applicant can get data on other sources expeditiously. One comment gave an example of the possible permutations and combinations involved in using multiple load levels for multiple sources, and suggested that, for a large city, the entire federal budget could not pay for modeling all the combinations. Another asserted that multiple load modeling for a variety of sources was not necessary, perhaps contributing an 11% increase in accuracy versus a basic inaccuracy of a factor of 2 in the models.

As the final topic under requriements for source emissions, the proposed EPA guideline addresses two aspects of growth. One commenter, a state agency, suggested the word "changes" because growth, in fact, might be negative. Two other commenters suggested that the guideline needs to provide more explicit guidance on this topic. A state agency comment objected to the inclusion of secondary and area source data because the emissions would be so poorly defined as to render any conclusions questionable.

Data Base For Background Air Quality

There were twenty comments on various aspects of determining "background." One commenter pointed out that the procedure required by the guideline now seemed to consist of a concentration estimate for the source in question plus modeled concentrations from other identified sources in the neighborhood plus non-specific background. He suggested that if this is the EPA requirement, it should be stated more explicitly in Section 2.2 of the guideline. As to the length of record of background data, another commenter recommended that a two-year period of monitored data be used to establish background, unless there are recent known emission changes.

EPA suggests in the case of areas with a few well-known sources but not urban areas, that background concentrations can be determined by using modeling to remove the effect of the known sources. Two state agencies commented on this procedure. One requested more guidance on how to remove the influence of existing sources. The other agency suggested that this procedure should not be required because the adjustment of background is subject to all the uncertainties of basic modeling. This agency suggested the use of upwind monitors instead. Another state agency requested further guidance on EPA's phrase "similar meterorlogical conditions." One commenter suggested that the procedures are not workable literally in some cases, while two other commenters suggested alternative procedures. Still another state agency indicated agreement with the procedure suggested in the proposed revised guideline for determining background by averaging monitor values when the winds are blowing in a way that the monitor is not impacted by the source in question.
A meteorological consultant suggested that periods when the

wind speed in less than five miles per hour be eliminated from the averaging when determining background.

There were several comments on the use of multi-source modeling to establish background. One state agency said that the guideline was unclear in regard to the determination of background concentration by monitoring or modeling in an urban area. The agency suggested that this section of the guideline should be rewritten to reflect a procedure consisting of modeling all point sources and adding a monitored value representative of the impact from all other sources in the area to the model predicted value, or else by modeling all point, area, and line sources in the area to determine impact. The comment of another consulting firm also agreed that multi-source models are needed to determine background in areas of high emission density. However, another firm suggested that one should not have to model small sources at a considerable distance from the source in question. The comment requested that EPA put forth a set of rules on this aspect. Another state agency suggested such a rule, which would be to require explicit modeling of point sources at distances shorter than five kilometers or perhaps to the distance of significant impact from the sources. One EPA regional meteorologist asked what should be done for urban sources farther than one to three kilometers from the source being modeled.

There were five comments dealing explicitly with determination of background suspended particulates. One consultant suggested that for determining background concentration of total suspended particulates, the usual procedure of excluding monitors within a 90-degree downwind sector is usually not applicable. He suggested multi-source modeling when the emissions are available. Two other comments simply stated that more guidance is needed in regard to determining background for total suspended particulates. Still snother comment indicated that the guideline should clarify whether the mean background for total suspended particulates should be an arithmetic mena or a geometric mean. Finally, another consultant suggested that there should be an adjustment made for differences in altitude when determining particulate background.

Placement of Receptors for a Regulatory Application

There were more than forty comments concerning section 5.1 of the proposed modeling guideline. One comment endorsed the concept of geometric progression of spacing between rings for the CRSTER model. One state agency participant explicitly supported EPA's recommendations for finding the distances to receptor rings for the CRSTER model. However, he suggested some modifications to the guideline guidance with regard to receptor spacing on rectangular grids. He recommended using a one kilometer grid spacing equal to the ninth ring distance for CRSTER, or to 10 kilometers whichever was less. He also suggested a grid spacing equal to one fifth of the downwind distance for consistency with the procedures for selecting ring distances. However, he recommended against using spacing less than one hundred meters or greater than one kilometer within the area so defined. One person suggested that the one kilometer spacing conflicted with informal guidance which had been received from EPA suggesting spacing as tight as one hundred meters. He recommended that the guideline explicitly state whether there are some situations in which a tighter grid is needed.

One comment stated that the recommended receptor grid spacing was too dense at distances greater than eight kilometers and too sparse to detect the highest concentrations close to the source. This comment suggested a variable grid spacing. The recommended grid spacing should be equal to the smaller of the radial or the azimuthal grid spacings for a polar grid, as would be determined by the use of PTPLU. This comment also suggested initially using a large-area grid, followed by additional model runs on successively finer grids in limited areas. Another comment indicated that the rectangular grid specification would be inadequate in the vicinity of large point sources and requested more guidance.

Approximately a dozen comments requested more flexibility in the placement of receptors. One of them indicated that 400 was about the right number but suggested that the placement be worked out on a case-by-case basis. Another comment stated that the strict adherence to the 10 receptor ring procedure was unwise. This comment also indicated that for rectangular grids the number of receptors could get very large without adding much to the amount of information. Many of the commenters who sought more flexibility echoed the previously expressed sentiment that the EPA procedures, especially the procedures for setting up rectangular grids, were inadequate. The rest of the comments on this particular topic only asked that more flexibility be allowed. One of these suggested that the guideline could set the minimum acceptable receptor requirements and additional refinements would be decided on a case-by-case basis. Another suggested that a one kilometer grid was too coarse for sources other than power plants, while another suggested that the polar grid procedures in the guideline are not adequate for low, relatively non-buoyant sources, such as baghouses.

As an aid to setting up receptor grids, three commenters suggested that the model PTPLU should be modified to include the option of using urban dispersion coefficients. One of these commenters indicated that neither the PTPLU procedure nor the one kilometer square grid are adequate for modeling downwash or fumigation.

Another group of comments suggested that EPA's procedures required too many receptors. One comment stated that ten rings for the CRSTER model is excessive, and two comments said that EPA should weigh the benefits against the cost before requiring that number of receptors. Another said such exhaustive arrays of receptors are wasteful.

One state agency suggested five kilometer spacing, with one-kilometer spacing in area of concern. Another state agency indicated that a one-kilometer spacing over an entire urban area (plus eight receptors around ever suspected hot spot) is simply too costly. The Ohio EPA cited an example of a county that would require 2,350 receptors according to the proposed guideline. Another state agency said that using 400 receptors is too costly and that extending a grid ten kilometers in every direction on a dense grid is financially impossible.

Several commenters discussed ways of reducing the number of receptors required. One suggested agreeing in advance that certain portions of a grid in an urban area would be insignificant and that receptors in such portions could be deleted. Others suggested a different approach to reducing the number of receptors. These suggestions were all based on using a short period of meteorology as a screening tool. One state agency suggested that the size and number of receptors be based on one year of modeling. A set of three substantially identical comments suggested the use of prescreening the meteorological data to identify persistent unfavorable conditions as a means of reducing the volume of data processed and the size of the receptor array. Another recommended approach was to screen the meteorological file for days of persistent unfavorable meteorological conditions and then to model only the worst days in the full grid.

One state agency indicated that the recommended spacings are not adequate for either polar grid or rectangular grid. Taking the opposite view, a set of comments objected to overly tight spacing, saying that one kilometer is the smallest meaningful separation. One comment said the guideline should be more explicit with regard to locating the additional receptors for hot spots.

Three comments addressed locations where receptors should not be placed. One suggested that no receptors should be placed over water; a second requested guidance on whether receptors should be placed on plant property, and a third said that receptors should not be required beyond fifty kilometers or the distance of de minimis impact, whichever is less.

There were three general comments on receptor placement. One indicated that the EPA procedures were inconsistent between the spacing of polar grid receptors and the spacing of rectangular grid receptors. Another said that EPA's instructions for locating receptors would not be applicable to the VALLEY model. The third comment relative to placing receptors for specific models suggested that EPA should state a policy on receptors less than 800 meters downwind from low sources when using the CDM model or the TCM model, claiming that the sector averaging procedure would produce excessive concentration estimates at shorter distances downwind.

Other Comments on Section 5

Other topics in section 5 include averaging time, treatment of wind data, pollutant half life, and classification of urban versus rural dispersion conditions. The comments on treatment of wind data and urban versus rural dispersions were summarized under Recommended Changes to EPA models.

The proposed guideline, in Section 5.2, gives the user a choice of either block averages or running averages. Two comments stated that the guideline should not give such a choice. Both of these comments suggested the use of block averages because that is the format in which most of the models provide output. Four other commenters also indicated a preference for the use of block averages. However, one of these suggested it might be appropriate to rerun the model to obtain running averages for the three-hour time period whenever the concentration estimated using block averages was larger than 80 percent of a standard or PSD increment. Still another commenter suggested using the block averaging procedure as a screening technique. However, above some threshold close to the size of available increment, he would use a running average. Two state agencies sugggested that all models should produce output compatible with EPA's interpretation of monitoring data, which is to use a running average. It was suggested that EPA modify their models accordingly. In a slightly different vein, one commenter suggested that all models should be built up from computations of hourly averages, which implies discarding such models as CDM and VALLEY.

The subject of pollutant half life has already been discussed to some extent in connection with the discussion of the VALLEY model. There were, however, five additional comments directed specifically at Section 5.4 of the guideline. One comment said that the guideline needs to provide more technical justification for the distinction made between the half life for sulfur dioxide recommended for the East and the half life recommended for the West. Two other commenters suggested that the recommended half lives are too short. One of them cited several references to studies showing longer half lives. One comment stated that the recommended cutoff at 50 kilometers for point sources in rural areas was arbitrary. He suggested sulfur dioxide conversion or removal be permitted at all distances even though the conversion or removal would be small at distances less than 50 kilometers. Finally, one state agency requested guidance on the appropriate half life for sulfur dioxide transported over water bodies.

The Use Of Measured Air Quality Data In Lieu Of Model Estimates

The final topic on which EPA had specifically requested comments was the use of measured air quality data in place of model estimates. One consulting firm expressed agreement with the EPA position as spelled out in the proposed guideline. Four industrial representatives and two state agencies suggested that modeling is not always more accurate than monitoring. The reasons given for disagreeing with the EPA position include: (1) meteorological data and emissions data are far from perfect: (2) ambient standards are very imprecise; and (3) modeling is now frequently done with such imperfections as using fictitious stack heights. One of the state agencies said that only monitoring can be used to determine actual air quality. Another comment said that monitoring should be encouraged as a check or substitute for modeling. One person indicated that EPA should consider that models are inherently inaccurate and that an on-site model is never fully validated until it has been "tuned" using monitoring data. Four comments indicated a particular circumstance in which monitoring is expected to be a valid indicator. That circumstance is the expansion of an existing source in an isolated area.

There were six comments on the general requirements in Section 7.1 of the proposed guideline. Four of these suggested that the criteria a through f were too demanding. In particular it was said that it was impossible to identify the absolute maximum concentrations with a finite grid. It was requested that the guidelines specify some acceptable minimum density of monitors. One state agency suggested that the monitor locations should have been based on an appropriate modeling. However, criterion f states that the recommended models must have been shown to be not applicable. One comment said that EPA needs to define when modeling is "not applicable." Another said, with regard to criteria e and f that, it is unfair to penalize a source for the vagaries of weather from one year to the next, and that the EPA models themselves have not been shown to be applicable.

There were two comments regarding the requirement stated in Section 7.2. One of them requested more flexible criteria, stating that complex terrain monitoring is so difficult that a more cost effective program is needed. The other comment stated that the 22½ degree sector described in criterion c of Section 7.2 might not be adequate.

Requests for Inclusion of Other Models

There were nineteen fairly independent requests for the retention of the Texas models, TCM-2 and TEM-8. Seven of these comments simply stated that these models should be retained. One suggested that these models be retained until EPA demonstrates that other models are actually more accurate. The remaining comments suggesting continued inclusion of these models pointed to the cost advantage these models enjoy compared with EPA models. The state of Indiana said, in essence, that the state should have the option to use an inexpensive model as long as there are no performance standards. Five comments stated that in addition to lower cost, the Texas models used more realistic sigma values. Two or three of these commenters also stated that some options of the Texas models could be made equivalent to the EPA models. One went so far as to recommend that the Texas models replace the EPA models, stating that the EPA decisions are arbitrary and that the Texas models include technical improvements recommended by the National Commission Air Quality.

Two other comments suggested that the model PAL should be retained also. The use of this model with settling was claimed to be an advantageous procedure for fugitive particulates. This comment also indicated that EPA should consider other models for fugitive particulate on a case-by-case basis.

Three participants suggested retaining the AQDM model. However, one of these recommendations was conditioned on EPA modifying the model to use the Briggs plume rise formula. The third expressed the same sentiment previously discussed in connection with TCM and TEM, that is, that the AQDM should be retained until EPA demonstrates that other models are actually more accurate.

Finally, there were requests for adoption of certain other specialized models. The Aluminum Company of America suggested that the BLP model should be adopted. The Department of Housing and Urban Development suggested that the guideline mention as screening models certain methods developed for that department by SRI.

Appendix E

MODELING LONG RANGE TRANSPORT AND DIFFUSION

Arthur Bass

Environmental Research & Technology, Inc.
Concord, Massachusetts 01742

1. INTRODUCTION

The growing concerns over the long range transport and transformation of sulfur oxides and other industrial effluents, and their subsequent consequences for human welfare and for the environment, are truly international. Studies to characterize and to model the regional and long range atmospheric transport, transformation, and removal of SO_2 have grown to become inter-regional and transfrontier in scope--as evidenced by, for example, the OECD program on the Long Range Transport of Air Pollutants (LRTAP). Closer to home, the recent establishment of a joint U.S.-Canada Research Consultation Group on the LRTAP problem, and the ongoing major initiatives by DOE/EPA in the Multistate Atmospheric Power Production Pollution Study (MAP3S) and by EPRI in the Sulfate Regional Experiment (SURE) to characterize and to model the regional formation and dispersion of sulfates (two of the many important observation/ modeling investigations of sulfur oxides transport at long ranges) clearly underscore the critical requirement for accurate means to predict and to verify the atmospheric dispersion of SO_2 over large spatial and temporal scales.

The title of this review is something of a misnomer, for it is hardly necessary to review for this audience the evolution of long range transport modeling. Indeed, several excellent, timely reviews abound: notable, for example, are the recent papers by Pack et al. (1978), Smith and Hunt (1978), Eliassen (1978), Fisher (1978), and the earlier work of Eliassen and Saltbones (1975), Fisher (1975), Scriven and Fisher (1975), among others, capture well the seminal results and the key conclusions obtained by the community of long range transport modelers, at least to the near present. It would be fatuous to attempt to distill here, in a few pages, the wealth of detail to be found in the recent conference literature on long range transport modeling studies: most notably, the Dubrovnik Symposium on Sulfur in the Atmosphere (Husar et al. (1978)), the Sofia WMO Symposium on the Long Range Transport of Pollutants (WMO (1979)), and numerous workshops and specialty conferences, as well as the sessions devoted to papers on this subject at the several AMS Symposia on Atmospheric Turbulence, Diffusion and Air Pollution (Santa Barbara 1974; Raleigh 1976; Reno 1979); and the AMS/APCA Joint Conferences on Applications of Air Pollution Meteorology (Salt Lake City 1977; New Orleans 1980).

The perspective taken in this review of long range transport models is much more limited in scope - for the most part, oriented to issues of concern for practical model applications to long range transport problems of pressing regulatory concern in the United States - that is, the transport of sulfur oxides and nitrogen oxides emitted principally from major point sources toward distant pristine areas (the so-called "Class I" Clean Air regions of the country); as well as the associated problem of acid precipitation impacts on regional and longer ranges, and that of preserving visual quality in Class I regions.

We do not, therefore, emphasize the several research-grade models in active development; but rather, we stress only currently operational, easily-obtained long range transport models. Specifically:

- models developed by U.S. research groups, that are already (or are soon to be) in the public domain;
- models that appear suitable to at least some present or anticipatable near-term regulatory applications;
- models that are practical, cost-effective, computationally manageable and highly user-oriented; or models that could be so made with minimal efforts.

This review tries to answer (if only in a preliminary way) the following questions, in keeping with the theme of this Conference - modeling in support of air quality regulations:

> Is the practical, routinely achievable state-of-the-art in long range transport modeling reflected in present regulatory practices or model guideline requirements?

> Are there impediments, in the present understanding of long range transport and dispersion, that militate against development of better verified, better justified long range transport models for regulatory use? What can be done, in the near term, to bring about useful improvements in long range transport modeling tools for practical regulatory applications?
>
> What prospects do the frontier areas of active research and development in long range transport modeling offer for improving predictive capabilities?

With this perspective, therefore, our review is "biased" towards models for long range transport and dispersion of weakly reactive pollutants—in practice, SO_2 - because the present state-of-the-art of photochemical dispersion modeling does not extend much beyond local or subregional scales (characteristic transport ranges of, say, 50 km or less). For purposes of practical applications, rather than for basic research or new modeling technique development, we will be concerned with models that include highly parameterized (essentially linear) plume chemistry. (Although important progress has been achieved in developing and verifying multispecies nonlinear plume chemistry mechanisms, these are not yet practicable for embedding within a mesoscale or long range transport model for routine regulatory use.)

Another major "bias" throughout this review is its focus on point-source-specific rather than area-source models--because by and large most regulatory decisions (and certainly the planning/site evaluation studies that precede a new source permit application) are concerned with incremental ambient pollution burdens that might be imposed by one or at most a small number of new sources. This is not true, of course, for regional-scale or national-scale planning studies (of the kind, say, that might be undertaken to support national energy policy on coal conversion, on interregional shipment of low sulfur coal, and the like); but the technical resources and modeling tools needed for studies of this scope are for reasons of economics and staffing limitations only available to the very largest research groups--the National Laboratories, EPA, EPRI, universities, and a few consulting organizations; the resources available in practice to a new source applicant, or a beleaguered local or state regulatory group, will usually be considerably less powerful and less sophisticated--and it is precisely these prospective users who have the most need for practical, credible, and regulatorily acceptable long range transport models.

A third major "bias", one that most underlies our later recommendations for model use, is toward models that can in principle be used to address short-term average ground-level concentrations (24-hour averages, and perhaps, 3-hour averages) at regional and long transport ranges. We emphasize here the short-term average models because the short-term PSD Class I SO_2 concentration increments to be protected under the 1977 Clean Air Act Amendments are usually more constraining to a new source applicant than are the annual average SO_2 increments. In view of the present and possible future regulatory needs for PSD and SIP review, we emphasize, and make recommendations concerning, models that are economical for evaluating the long range impacts of one or at most several sources (on the order of ten, say), rather than large grid-type multisource models best suited for regional type inventories.

We will show a consistent bias toward models that are easy to use and easy to explain to the nonspecialist:

- models that are suitable for a broad range of meteorological dispersion conditions - both worst-case (episodic) and average dispersion;
- models that require only readily obtainable meteorological and other input data;
- models that are flexible, modular, readily transported and modified for use on computer systems of conventional size;
- models that are well documented for use outside the originating organization;
- models designed from the beginning for efficient, economical simulations, if possible, both in the near field of individual sources, and in the far field where combined effects of multiple sources may dominate.

Finally, we have tried to emphasize models that recognize explicitly the physical processes that govern turbulent plume transport on regional and synoptic scales--but not models that, in our subjective opinion, overemphasize secondary parts of the modeling problem, while ignoring or giving short shrift to the primary, crucial issue: realistic description of mesoscale and long range turbulent transport in a spatially and temporally varying mixed layer driven by synoptic-scale weather systems.

For purposes of discussion, we will sometimes use the term "long-range" loosely, to mean transport distances beyond about 50 km, notwithstanding the important physical distinctions between transport on the β-mesoscale range (on the order of 25-250 km) and transport in the α-mesoscale range (on the order of 250-2500 km)—see Gillani (1978) for further discussion. Both pose important problems for regulatory decisionmaking: the problem of significant near-field ground-level concentrations of SO_2 falls within the short range end of the β-mesoscale; whereas the problem of ambient sulfate impacts, including issues of acid precipitation and regional visibility degradation, extends well into the α-mesoscale.

It is important to bear in mind, therefore, the current regulatory position on long range transport - as stated in the EPA Guideline on Air Quality Models (EPA 1978), estimation of source impacts is highly uncertain beyond 50 km: the choices of dispersion coefficients are "tenuous," and diurnal variations in meteorology over such transport times "are more likely to alter plume trajectories and dispersion characteristics." The Guideline asserts that no widely applicable model is available for dealing with long range transport and removal - and that source impacts at large transport distances "should be considered on a case-by-case basis with available techniques." This paper attempts to focus attention, therefore, on the full range of techniques available now for practical use.

As basis for the suggestions we will subsequently make, we illustrate several generic approaches to Lagrangian modeling on regional and longer scales; describe some of the many possible algorithm choices adopted by different research groups; briefly survey the available model sensitivity and model verification studies for selected models; review the requirements for Lagrangian long range transport models in light of current and possibly near term ambient air quality and air quality-related regulations; and then list some available long range transport models (principally Lagrangian models but also some numerical grid models) that are currently operational and already (or soon to be) in the public domain, so far as we know.

Finally, we identify some of the current directions of active research and development in long range transport modeling, and suggest some of the future directions in which such modeling efforts might be directed for achieving practical model development objectives. Other papers to be presented at this Conference will describe more fully some of these principal directions in LRTAP model development and verification at the National Laboratories and elsewhere--efforts that seek to extend the present understanding of the physics of regional and long range dispersion, and to develop newer and more powerful computation methods to address these problems.

2. CURRENT REGULATORY NEEDS FOR LONG RANGE TRANSPORT MODELS

At present, reliable, practical long range transport modeling tools are urgently required, both for regulatory applications and for other planning needs for example:

- for development and revision of SIP-related regional control strategies;
- for New Source Review programs to attain and maintain National Ambient Air Quality Standards (NAAQS), and for the Prevention of Significant Deterioration (PSD) in Clean Air regions; and
- for a wide range of policy and planning studies to assess the large-scale, long-range environmental, economic, and social implications of new coal-based energy resource development, fuel switching and coal conversion orders, delays and/or cancellation of nuclear and hydroelectric power generation capacity, and like issues.

The modeling techniques specifically referenced in the Guideline--Heffter and Ferber et al. (1975); Rao et al. (1976); Scriven and Fisher (1975); and Hales et al. (1976)--have been supplanted by more powerful and more useful models, better able to address directly the specific spatial and temporal resolution and averaging requirements imposed by the current ambient air quality regulations.

To meet these ambient air quality source review requirements, the more urgent need now is for credible, well-verified short-term average (episode-type) models; long-term average models are of lesser priority at present. However, where the regulatory decisionmaking process must deal with secondary pollutant effects such as long-term average acidification and biological impacts, the need for climatological models on regional and longer spatial ranges will be much greater. As driven by present regulatory practice, the near-term requirement is for models of spatial resolution and domains of validity appropriate to the β-mesoscale (for a typical β-mesoscale impact problem--the maximum ground-level concentration of SO_2 from an elevated point source at a range on the order of 100 km); other models, or perhaps the same models with resolution more suitable to the α-mesoscale range, would be best for treating, say, the problem of modeling maximum ground-level concentrations of sulfate formed by conversion of SO_2 from elevated plumes ducted aloft for long distances.

As a rule, the computational approach and resolution of a model can be economically matched only to problems of a certain spatial/temporal range. For example, the "square puff" trajectory model approach described later is extremely efficient for long-term average dispersion over large distances; the same model, however, will underpredict drastically in the near field of a source. As another example, a long-range transport model that requires immediate uniform vertical mixing of an effluent plume, from the outset, may be entirely suitable if the concern is only about impacts over travel times comparable to the diurnal cycle of the mixed layer, but would entirely misrepresent the ground-level impacts in the near field of an elevated point source under stable nighttime flow.

To meet the short-term average requirements, the needed temporal resolution of long-range models for regulatory application is on the order of one to a few hours. Yet, because it is necessary to predict maximum 3-hour average or maximum 24-hour average ground-level concentrations not to be exceeded more than once per year, models that can economically simulate one year or more of 3-hourly average concentrations are also needed now.

Among other regulatory applications of long range transport models, it is possible that modeling of long-range visibility impairment will play an important future role in PSD source review, especially in the western United States. Visibility degradation is both a β-mesoscale and an α-mesoscale problem.

Plume blight and plume discoloration are related chiefly to ambient concentrations of suspended particulates and NO_2 within, say, 100 km of a typical new major coal-burning source. To model visibility impacts from individual sources economically on these transport scales, Lagrangian transport models coupled to radiative transfer algorithms will be needed; several such models are already available. On longer transport scales, say at a distance of several hundred kilometers, the visibility impairment problem is one of regional haze; and successful, credible simulations of multisource impacts on regional visibility at such scales are not likely to be achieved very soon.

We may also anticipate an evolving requirement for the use of long range transport models to address regulatory needs for modeling acid precipitation impacts. In contrast to visibility impairment problems, which are related to instantaneous ambient concentrations of species causing light extinction, acid precipitation problems include both episodic and long-term average concentrations.

Greater emphasis can be expected on policy/planning and regulatory applications of long range transport models for inter-regional transport and diffusion across state boundaries, as for example, the growing controversy over SO_2 emission levels in the Ohio River Valley, or the increasing concern over transnational boundary fluxes of SO_2 and sulfates between the United States and Canada.

Sometime in the future, large numerical grid models for long range transport will play a major (perhaps dominant?) role in practicable regulatory decisionmaking; but at present, and probably in the near term, for most applications Lagrangian long range transport models will continue to be more suitable—for reasons of near-field resolution, data requirements, computer resources, costs, and ease of use--at least for single sources, or for small ensembles of sources for which the plume chemistry is sensibly linear.

3. LAGRANGIAN MODELS FOR LONG RANGE TRANSPORT

A Lagrangian variable trajectory plume model represents a continuous plume emitted by a point source by the transport and dispersion of a succession of discrete plume elements (air parcels or massless trajectory points). These plume elements or air parcels are independently advected and diffused by a spatially and temporally varying wind field. Each plume element carries an independent time history--plume chemistry, dry deposition and scavenging. The time-average ground-level impact of the (continuous) plume at a given point is simulated by combining the contributions from all elements that independently traverse that point during the specified averaging time. Lagrangian models thus offer the opportunity to resolve the near-field small-scale impacts of individual point sources with resolution that grid models cannot achieve, at least not economically. (For this reason, some grid models (e.g., Liu and Wojcik 1979) include subgrid scale plume or puff-type modules to carry the early evolution of a plume element to the time when the element has grown sufficiently large to be "handed over" to a grid cell, thus preventing unacceptably large initial dilution.)

Lagrangian models fall into two broad classes. The nonreactive models employ simple linear or pseudo-first-order chemistry; handle multiple sources by superposition; and usually have a detailed treatment of horizontal plume transport and diffusion but generally little vertical resolution. Nonreactive Lagrangian models usually attempt to recreate in detail the mesoscale meteorological variations that dominate plume dispersion on these transport scales. These models tend to be relatively inexpensive per source simulated.

Reactive Lagrangian models, by contrast, include complex multispecies chemical mechanisms; use a single "reactor" volume advected by the local wind; and, typically, have only very limited horizontal coverage: essentially two-dimensional along the trajectory. The vertical resolution is much more detailed, because the plume chemistry will typically vary strongly with altitude. Reactive Lagrangian models use, as a rule, only simple treatment of plume diffusion.

Nonreactive Lagrangian models (the only ones to be discussed here) can be further partitioned into short-term average and long-term average models. Short-term average Lagrangian models attempt to resolve individual plume trajectories using sequential meteorology. Because the short-term models include so much detail about the transport, diffusion, and transformation of many individual plume elements, some may be computationally limited to handling a small number of individual sources.

Some long-term average Lagrangian models include only a simple treatment of horizontal diffusion or may ignore horizontal diffusion altogether, as does the PNL regional model (Wendell et al. 1976). Other long-term average Lagrangian models have described both horizontal and vertical dispersion processes, each in great detail, but as completely decoupled prcblems; for example, Bolin and Persson's (1975) model or the ASTRAP model (Shannon 1979a). Still other long-term average models may employ simple assumptions on the vertical distribution of plume material—for example, the uniform mixing assumption of the EURMAP-1 model (Johnson et al. 1978). These models are designed, foremost, to be computationally economical for one or more sources over long periods of time; they cannot therefore also provide the near-field detailed spatial resolution that may be needed for regulatory evaluations.

Most Lagrangian plume models begin from the premise that long range transport and diffusion (strictly, for transport over times longer than a few averaging times) is dominated by plume meander due to mesoscale variations in the wind field—that is, by large eddies of characteristic scale greater than the plume dimension—what Powell (see Nappo 1978) has called "mesoscale turbulence"--rather than by small scale turbulent eddies smaller than the plume size. Figure 1 illustrates schematically three different regional-scale Lagrangian trajectory modeling approaches. (Other Lagrangian models, such as Sheih's (1977)

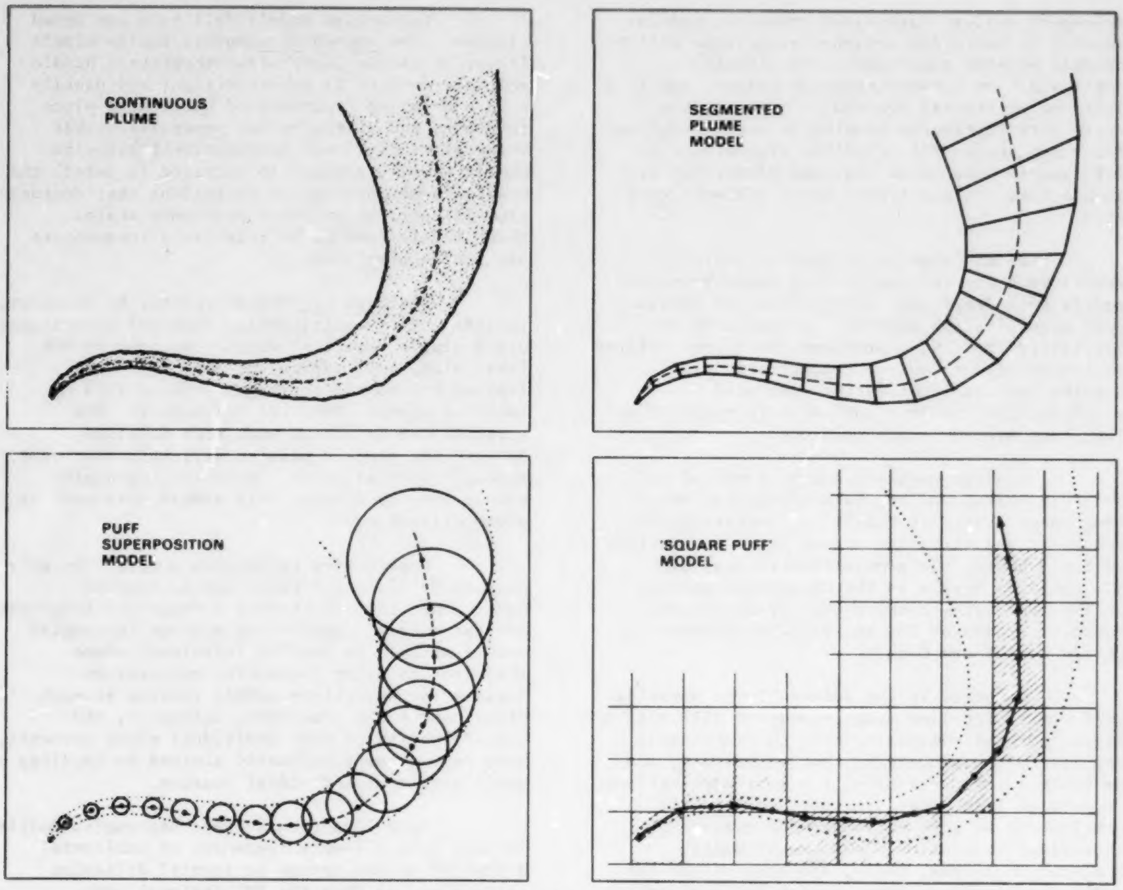


Figure 1 Lagrangian Trajectory Modeling Approaches

statistical trajectory approach or his six-particle puff model with wind shear (Sheih 1978a), are less readily visualized).

The puff superposition model represents a continuous plume as a series of discrete puff elements of circular horizontal cross section; the vertical mass concentration field may be treated as Gaussian, as uniformly mixed, or as prescribed by a finite-difference representation of the vertical diffusion equation (see for example Draxler 1979). The puff release rate and puff advection times must be chosen in a consistent manner, so that consecutive puffs will overlap sufficiently to ensure a reasonable representation to the continuous plume. Ludwig (see Nappo 1978) provides guidance for the appropriate puff spacing to achieve coverage comparable to the continuous plume.

The segmented plume approach shown in Figure 1 differs from the conventional Gaussian plume model in that the segmented plume may be continuously deformed by a temporally varying horizontal wind field. The plume is treated as divided into contiguous segments. Each segment describes a portion of plume behavior between succesive time periods, and the end points of each segment are advected in a Lagrangian sense.

Under conditions of uniform, homogenous flow, the segmented plume and puff superposition models will give about the same results, and the plume model is typically about 50% cheaper to run. Why then are puff models favored by many workers? This issue was strongly debated at the 1977 Trajectory Modeling Workshop (Nappo 1978)--debated, but not entirely resolved.

The preference for the puff model derives not so much from theoretical considerations as from the practical advantages of the puff model: as expressed by Sheih (see Nappo 1978) puff models are generally more credible and less subjective to pathological behavior, especially under conditions of recirculating or strongly sheared wind flows, conditions for which a plume model may offer poor coverages at the interfaces where plume segments are "attached". And, especially under stagnating flow conditions, the plume model has an explicit inverse-wind speed singularity; when the winds approach calm, the model can violate the basic assumption of the plume approach--that mean advection is greater than turbulent diffusion in the along-wind direction.

Although the conceptual basis for puff superposition and plume segment models was not fully resolved at the Workshop--and remains unresolved at present--the rejoinder by proponents of these modeling approaches is that (a) they make better sense physically under irregular flow situations than the conventional straight-line Gaussian plume model they are meant to replace, and that (b) in practice they work better for regional scale and long range flows.

A central issue that must be confronted in using these variable trajectory models is the problem of how to partition the turbulent spreading into (a) large scale meandering of the plume centerline and (b) small scale turbulent spreading relative to the centerline, that is, how to partition the total crosswind spread. In practice, most modeling groups make a somewhat arbitrary choice of averaging time to describe the mean wind field (for example, one hour average winds). This mean wind field is used to advect the centroid of the puff or plume element. Lateral dispersion relative to the instantaneous position of the centroid is then prescribed as a function of travel time or downwind distance. For example, Start and Wendell (1974) use Yansky's sigma-curves; Heffter et al. (1975) permits the horizontal puff radius to grow linearly with time; Johnson et al. (1978) in the EUROMAP-2 model, allow σ_h to depend on the local deformation field; and finally, the ERT group, to make their puff model more consistent with routine regulatory usage, assume that σ_h evolves as described by the conventional PGT curves to a distance of 100 kilometers, and thereafter assume (with Heffter) that puffs grow as simple linear functions of time.

It is important to recognize that there is no rigorous theoretical justification for the artificial decomposition of turbulent dispersion into mean and small scale relative diffusion; and certainly not for the use of PGT or other conventional turbulence typing schemes based on the continuous straight line plume. But again, as expressed by several participants in the Trajectory Modeling Workshop, these puff and segmented plumes schemes, however artificial their decomposition of mean and relative diffusion, are useful.

The third kind of Lagrangian trajectory model shown in Figure 1 does not appear to have a standard name--so, for want of one more descriptive, we have called it the "square puff" model or the "massless point in box" approach. This scheme--a computationally efficient, long-term model developed by Powell and co-workers at PNL--is conceptually a descendent of the Start-Wendell MESODIF puff model (Start and Wendell 1974) except that only the puff trajectory remains; the horizontal diffusion of puffs is always fixed at the size of one grid cell. Powell et al. (1979) argue that for long-term averages at large transport differences, horizontal diffusion can be ignored; for if the trajectories are not systematically biased, then on the long-term average adjacent cells will see comparable diffusive contributions, and the net effect on a regional average basis will be about the same as if actual lateral dispersion were included. This is indeed an attractive, efficient approach, if the concern is not with near-field impacts (for which the model must necessarily underpredict the maximum ground-level concentrations, and overpredict virtually everywhere else), nor where the emphasis is on short-term predictions (where the lateral diffusive contributions of individual puffs are all-important).

Finally, Sheih (1977a) and Shannon (1979a) use a large ensemble of individual trajectory calculations to develop empirical equivalent puff dispersion statistics for their long-term long range statistical puff model.

Just as many approaches have been explored for characterization of the horizontal transport and diffusion of plumes over long transport ranges, so too have numerous schemes been suggested to characterize vertical diffusion. Several Lagrangian models describe vertical diffusion by a Gaussian profile - where the vertical plume spread statistic, σ_z, is specified in the manner of the conventional PGT σ_z curves, or else is specified as a function of vertical eddy diffusivity K_z and of travel time. Lagrangian models that specify σ_z in a manner similar to the PGT curves (which corresponds most closely to current regulatory useage of straight line model at shorter ranges) include for example: STRAM (Hales et al 1977), EURMAP-2 (Mancuso et al. 1979), and MESOPUFF (Bass et al. 1979), among others. By comparison, the ARL-ATAD model (Heffter 1980) specifies σ_z by a Fickian diffusion law. Many of these models also permit, as an option (if not as a fixed requirement) that the vertical Gaussian profile be replaced by uniform vertical mixing, either from the outset or when the puff/plume element has grown to appreciable vertical extent. Finally, several more recent Lagrangian puff type models use K-theoretic treatments of vertical diffusion by numerically differencing and solving the vertical diffusion equation as a multilevel problem including sources and sinks. Example models include the NOAA/ARL Mesoscale Trajectory and Diffusion model (Draxler 1979) and the BNL AIRSOX model (Meyers et al. 1979).

Shannon (1979b) has also used a Gaussian moment-conservation method for vertical diffusion in the ASTRAP model--this is a generalized Gaussian puff scheme, consisting of Lagrangian advection (countergradiently) by a fictitious diffusion velocity, followed by Eulerian decomposition and mass and moment conserving interpolation back to the Eulerian grid.

4. NUMERICAL GRID MODELS FOR LONG RANGE TRANSPORT

A detailed discussion of grid modeling approaches to regional scale transport is well beyond the scope of this review, but detailed descriptions of these modeling approaches and their verification histories are amply documented elsewhere; indeed, important work, only recently emerging, will be reported at this Conference by the EPRI/ERT, MAP3S, and other major research teams.

We should note that the ongoing development of the EPRI Sulfate Regional

Experiment and its associated model refinement/verification activities (the SURAD model) is discussed by Lavery et al. in these proceedings. Other important grid model development initiatives under the MAP3S program are being actively pursued by the BNL group (P. Michaels, personal communication) - notably, SCHEMATIC (MAP3S Modeling Abstracts, in preparation). A regional scale grid model using the pseudospectral method of Prahm has been tested against the SURE intensive data base (Niemann et al. 1979) with some success.

It should also be noted that the EPA Meteorology Laboratory has undertaken an ambitious multiyear research and development program to develop a "regional super model" including reactive chemistry, spatially and temporally varying mesoscale meteorology, and a rigorous treatment of turbulent dispersion on local and regional scales (K. Demerjian, personal communication).

But, for practical applications to regulatory problems, in the near term at least, we would judge that only the Northern Great Plains Regional Model (Liu and Durran 1977) and the MESOGRID model (derived from the EPRI/SURE regional model SULFA3D - Egan et al. 1976, Morris et al. 1979) are currently available.

5. METEOROLOGICAL DATA REQUIREMENTS FOR LONG RANGE TRANSPORT MODELING

5.1 Wind Field Synthesis

The experience of many groups suggests that the predictive success of long range transport models is affected most by the representativeness and adequacy of the available mesoscale-synoptic meteorological data base. Judging by the many published model sensitivity analyses, and the as yet fragmentary but rapidly improving efforts at verification of long range transport models, it appears that the principal present limitation to greater accuracy of these models is insufficient spatial and temporal resolution of the wind fields.

The effective spatial resolution of the upper level wind field observations is set by the characteristic distance between rawinsonde stations. The spatial resolution of the upper air sounding program in the U.S. is of order 300-500 kilometers; marginally adequate at best even under ideal conditions of uniform topography and reasonably homogenous flows. The temporal resolution of the upper air sounding program is a yet more serious limitation--soundings are only made at 12 hour intervals, and important diurnal variations in wind field, mixing height, and turbulent structure occur on much smaller time scales.

One possible approach used to transcend the limited spatial resolution of the upper level mesoscale wind data is that of Heffter and co-workers (Heffter et al. 1975, Draxler 1977, Draxler 1979); in their approach, the twice-daily upper air data is supplemented by the much denser set of hourly surface station data, interpolated upwards to represent the layer-averaged wind fields. This approach is not likely, of course, to be successful in regions of strong topographic influences or under severe convective conditions that may dominate mesoscale flow fields--for example, in the southwestern United States.

As another possibility, a mesoscale Numerical Weather Prediction (NWP) model of demonstrated prognostic skill in the geographic region of interest (over a period of, say, 12 to 18 hours) could be used to drive a trajectory model. This is the approach of the Drexel-NCR-BNL Limited Area Mesoscale Prediction System (LAMPS)--a system designed to simulate the evolution of mesoscale weather systems for periods of 24 to 36 hours, with grid resolution of 35 to 140 kilometers (MAP3S Modeling Abstracts, in preparation). (Preliminary forms of this modeling system have been exercised at BNL and additional work is planned to couple the mesoscale prediction system to a trajectory model.) Indeed, looking to the future, such combined mesoscale weather prediction-plume transport models may offer a practical way in which to supplement sparse meteorological data on mesoscales. For long range transport modeling applications in the near future, however, one is constrained to "make due" with the available observational network.

We have stressed the controlling influence of mesoscale meteorological flow conditions over long range plume transport and dispersion; and a brief, graphic example is perhaps appropriate here. This example is taken from a recent simulation study of regional scale air pollution associated with coal-based energy resource development in the Four Corners region of the southwestern United States (Bass et al. 1979). One of the principal study conclusions was that plume dispersion in the Four Corners region appears to be dominated by large scale variations in mesoscale wind fields, and by the spatial and temporal variations in mixing heights and stability class--but not, as a rule, by small scale dispersion. (This point has been made repeatedly, in the past, by the SRI, NOAA/ARL, and Battelle PNL modeling groups, as well as in the earlier work of the OECD LRTAP investigators, among others.)

Figure 2 and 3 illustrates the surface weather map and upper air flow fields for a typical winter "best-day" regional dispersion situation in the southwestern United States. The 700 mb winds were generated from the regional rawinsonde network data by the MESOPAC mesoscale meteorological preprocessor (Benkley and Bass 1979c). The westerly upper level flows are rapid and zonal (winds up to 25 m/s); the surface (gradient) winds are also westerly. Computed mixing depths (Benkley and Schulman 1979) were about 1,000 meters, caused by strong mechanical forcing of the surface boundary layer with little diurnal variation.

The 24-hour average ground level concentrations of SO_2 and SO_4 corresponding to this best-day winter dispersion case in the Four Corners area are shown in Figure 4. The SO_2 plumes, clearly disjoint, are advected strongly toward the eastern grid boundary, with rapid downwind decrease of concentration. The

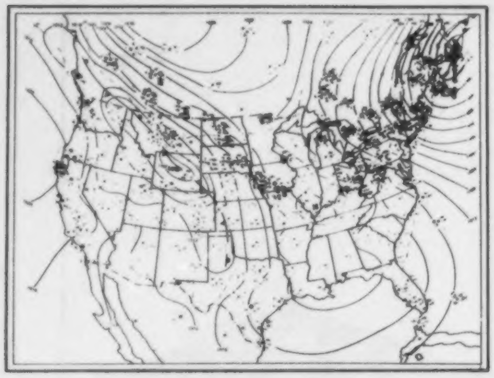


Figure 2 Surface Chart, 1200 GMT, December 7, 1977

regional-scale ground-level SO_4 concentrations are negligible--the maximum sulfate concentrations, on the order of only 0.1-0.5 $\mu g/m^3$ appear only in the downwind portions of individual plumes.

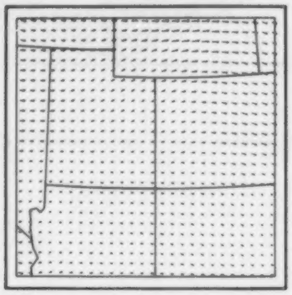
This best-case situation may be contrasted with the next example--'worst-case' dispersion. Figures 5 and 6 show the surface and upper level meteorological flows for a typical "worst-case" regional dispersion situation in the Four Corners area--a stagnant summer day characterized by weak surface pressure patterns and a surface heat low over the southwestern portion of the region. The upper level flows at 700 and 500 millibars are very weak and disorganized (wind field less than 2 m/s), with no generally distinguishable transport direction. The computed mixing depths are low during the night (on the order of 100-400 meters), and high during the day (on the order of 2,000-3,000 meters). Plume elements more than one day old are well mixed through a deep layer. Evidently, deep mixing does not preclude large sulfate accumulations on regional scales.

The 24-hour average ground-level concentrations of SO_2 and SO_4 for this worst-case summer day are are shown in Figure 7. They differ strikingly from the previous example: the SO_2 field shows considerable merging of individual plumes, and accumulation in the north central portion of the computational region. The sulfate field shows dramatic accumulation in the north central portion of the grid, with maximum concentrations greater than 10 $\mu g/m^3$.

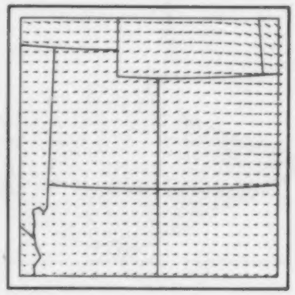
In all, 106 days during 1977 were modeled; these shown here were the best-day and worst-day regional dispersion regimes. The success of a long range transport model is, these results suggest, largely dependent on the physical verisimilitude of the meteorological "package" used to drive the models--among the least well understood parts of the long range transport modeling problem.

It is not surprising, therefore, that the available models differ appreciably in the way they define and use the meteorological information. The models differ especially in the definition of the advective wind field, that is, in the choice of representative level, in the scheme used to analyze and interpolate wind data, in the time averaging assumptions, and in the dynamical or kinematic constraints, if any, imposed in constructing the mean flow field.

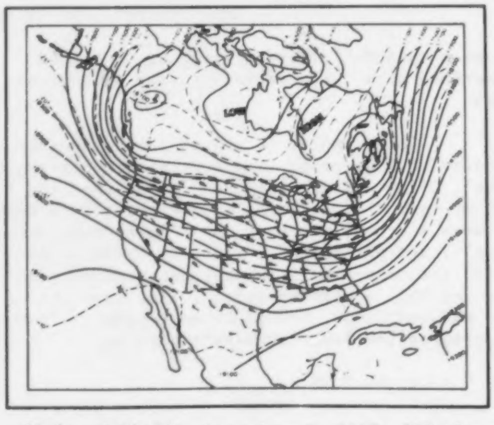
The central issue is: How is the mean wind field to be specified? Lagrangian trajectory models require a single transporting wind field, assumed to be representative (in some sense) of mean transport layer flow conditions over the scale of interest. (Some vertically averaged grid models--like the SAI model--also use a vertically averaged mean wind field.) The wind field may be that at a single level (for example at a constant height above the surface),



(3a) 1200 GMT, December 7, 1977, 700 mb

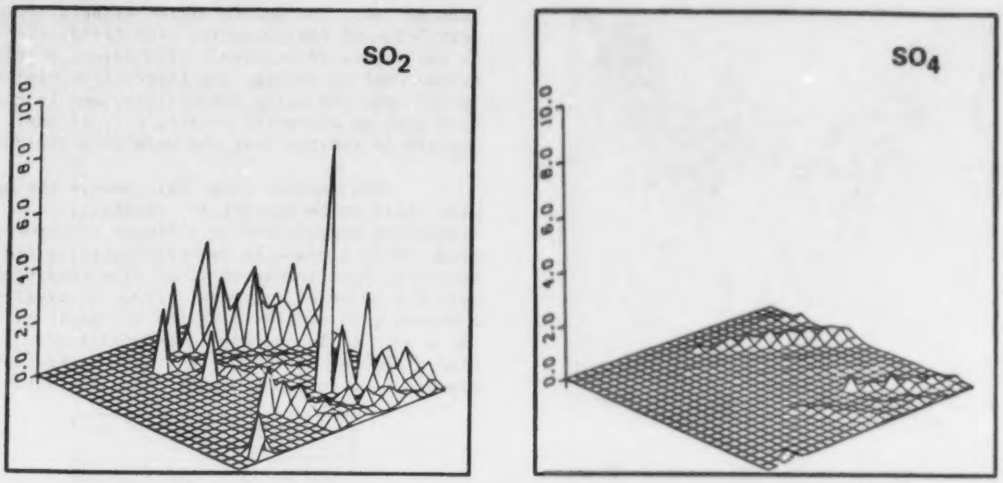


(3b) 00 GMT, December 8, 1977, 700 mb

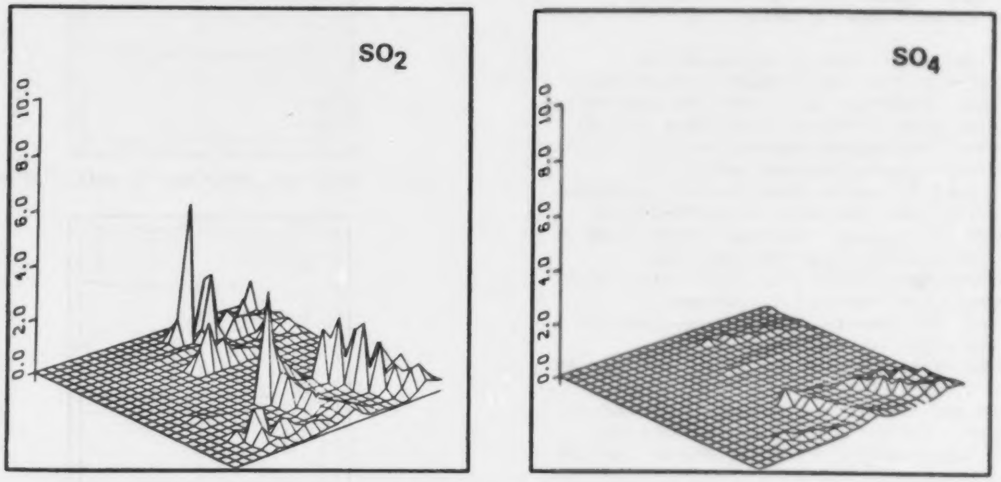


(3c) 1200 GMT, December 7, 1977, 500 mb

Figure 3 Upper Air Meteorology



DECEMBER 6, 1977



DECEMBER 7, 1977

Figure 4 "Best Case" Regional Dispersions, Four Corners Area (Winter Zone Regime). Daily Average Ground-Level Concentrations ($\mu g/m^3$)

or on a constant pressure surface, or may be described as a layer average in any of several ways.

For example, in one of the forerunner studies of long range transport with a Lagrangian model, Eliassen and Saltbones (1975) used the 850 mb wind field, linearly interpolated in time between six-hour observations. Among more recent models, the SRI EURMAP-1 model (Johnson et al. 1978) for long-term interregional transport of SO_2 and SO_4 in western Europe, also uses the 850 mb wind field, but adjusted to give a "more representative layer average"; that is,

$$\bar{U} = 0.75\, U_{850mb}$$

$$\bar{\theta} = \theta_{850mb} - 15^o$$

The SRI short-term model, EURMAP-2, (Mancuso et al. 1979) requires finer resolution. It uses a weighted average of winds in two layers, 0-300 meters and 300-1,000 meters; within each layer the layer-average winds are obtained by integrating with a power law of the form

$$\left(\frac{U_2}{U_1}\right) = \left(\frac{Z_2}{Z_1}\right)^{P(\text{stability})}$$

In areas dominated by maritime meteorology, (e.g., Scandinavia, the United Kingdom, and other parts of western Europe), the trajectory errors made when 850 mb winds (or, for that matter, surface geostrophic winds) are used may be tolerable (Smith and Hunt 1978). By contrast, under more continental mesoscale flow

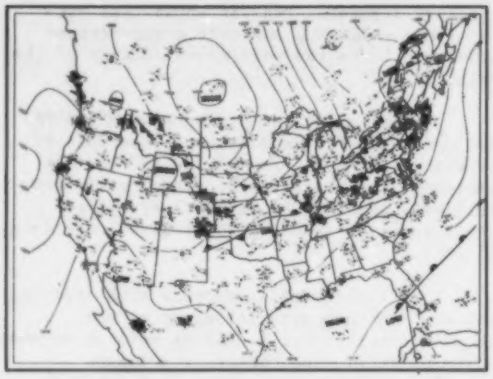


Figure 5 Surface Chart, 1200 GMT, July 24, 1977

regimes (especially in the winter over the U.S. Northern Great Plains, for example), the atmosphere is strongly stratified, and the surface and 850 mb winds are uncoupled. The 850 mb wind field (or in moderate to elevated terrain, the 700 mb winds) might be a more appropriate choice to describe mean transport conditions. But without a detailed verification study, like that made by Draxler (1979) on the sensitivity of his model to different assumptions on wind speed and wind veering with height, it is impossible to "know" which assumptions are best, particularly for a new application in a region for which adequate mesoscale meteorological data coverage does not exist.

The possible schemes for constructing mean wind fields are endless--which is unfortunate, because there has been little systematic attention given to model intercomparisons with consistently defined wind fields (but cf. Bass et al. 1979).

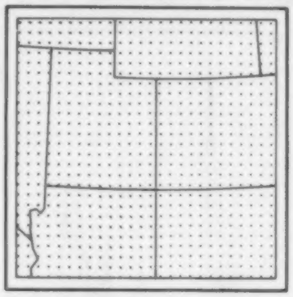
Each modeling group adopts its own unique scheme (or schemes). For example, in the northern Great Plains regional modeling (Liu and Durran 1977; Liu and Wojcik 1979), the 850 mb geopotential height field, spline interpolated, was used to generate a geostrophic wind that was then iteratively relaxed.

In the PNL regional scale model Wendell et al. (1977) used a layer average wind field averaged through a fixed layer (100-1,000 meters). They began from 12-hourly NWC rawinsonde station data, interpolated linearly in time, and used an inverse square spatial weighting to interpolate the discrete data points to a uniform grid. (The subject of interpolation schemes might well take an entire review paper; Goodin et al. 1979 have made an excellent summary of interpolation methods for sparse data applied to wind and concentration fields-to which the reader is strongly recommended.)

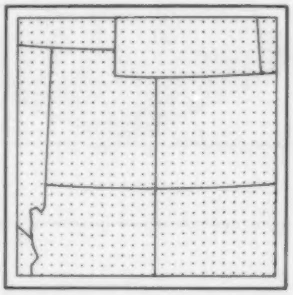
Among the other alternative approaches to wind field definition, Heffter and co-workers (Heffter 1980) define a spatially and temporally variable layer for averaging the wind field, based on the height of any nonsurface-based inversion as obtained from vertical temperature profiles. By contrast, Heffter's earlier long-term model used a single fixed layer for averaging, 0-1,000 meters.

Draxler's mesoscale model (Draxler 1979) carries Heffter's approach further--he averages the winds through a vertical layer of variable thickness--determined as the vertical extent within which 90% of the columnar vertical mass distribution is found. This layer is computed from numerical solutions to the vertical diffusion equation with specified vertical eddy diffusivity profiles.

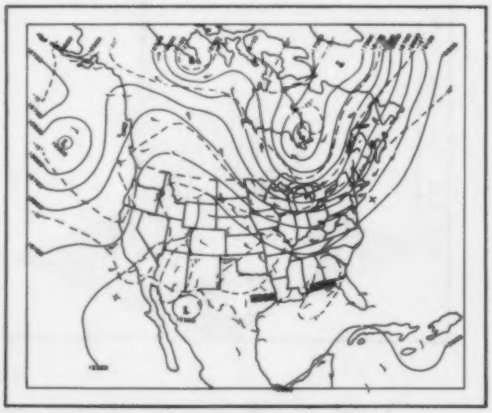
In both the Heffter and the Draxler schemes, the winds are averaged with a weighting that depends on the thickness between mid-points of observation levels - the surface wind data when used is adjusted upwards to represent the layer-averaged winds, and grid point interpolations are made within a radius of



(6a) 1200 GMT, July 24, 1977, 700 mb



(6b) 00 GMT, July 25, 1977, 700 mb



(6c) 1200 GMT, July 24, 1977, 500 mb

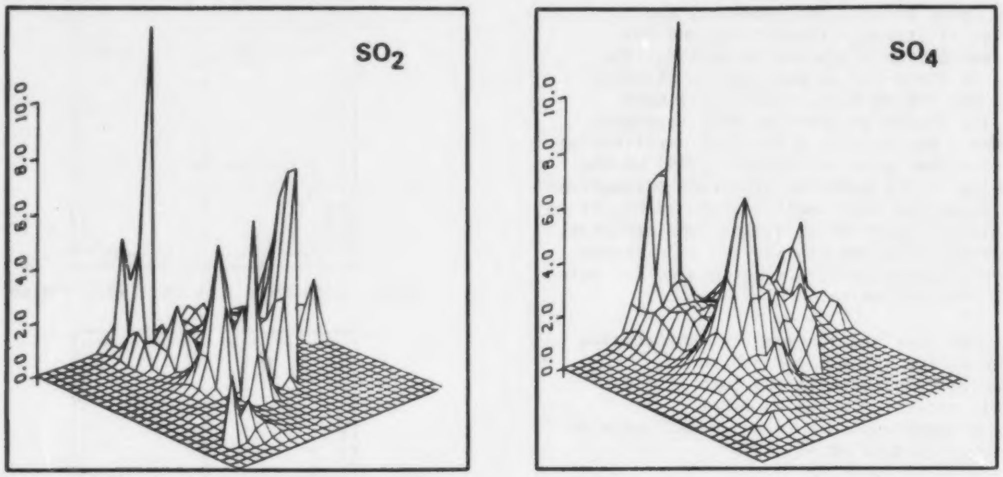
Figure 6 Upper Air Meteorology

influence on the order of 300 km for upper air stations, and 150 km for surface stations.

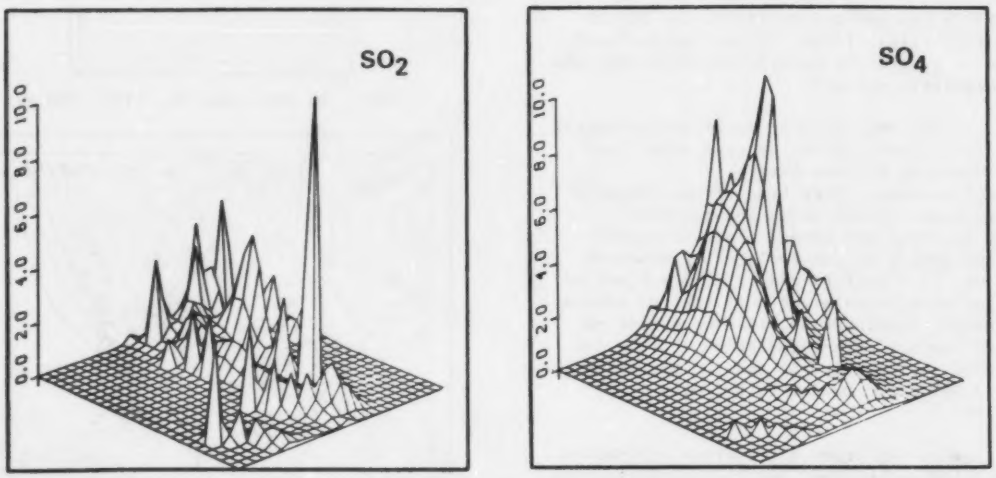
Not surprisingly, each modeling group has its own ideas about appropriate 'regions of influence' for station data interpolation, and about relative weights to be assigned to observation stations in performing the wind interpolations. Clark and Eskridge (1977) have made a varition on the Liu-Goodin scheme (Liu and Goodin 1976) that allows the user to specify a reliability or confidence factor for each observation station and to construct the interpolated wind fields by a composite of these weighted observations. The NOAA/ARL group carry the 'region of influence' concept one step beyond: their models (for example, ARL-ATAD) also give greater weight to observational data from stations that are aligned within a narrow angular range of the direction of the local trajectory segment. Draxler shows that such directional weighting can make a significant improvement in to the predictive success of the model.

Finally, ERT's multipurpose MESOPUFF model uses wind fields developed from data at user-specified levels (nominally 850 mb, for example)--interpolated to a uniform grid--and iteratively adjusted to within a specified maximum value of absolute local divergence (e.g., 10^{-4} s^{-1}).

In regions of widespread high terrain, the Rocky Mountain area for example, and especially where terrain extends well up to about the 700 mb level, the choice of a 'most representative' mean advecting wind field is extremely problematic at present (and likely to remain so). The conventional upper air



JULY 24, 1977



JULY 25, 1977

Figure 7 "Worst-Case" Regional Dispersion, Four Corners Area (Summer Stagnation Regime). Daily Average Ground-Level Concentrations ($\mu g/m^3$)

rawinsonde network cannot begin to resolve terrain-induced spatial variations in the upper level wind field, the surface station data is often useless, and local (e.g., source applicant sponsored) sounding programs can only indicate initial plume direction. In modeling the regional impacts of energy resource development in the high terrain flow regimes of the Four Corners region, Bass et al. (1979a) chose therefore to represent the mean transport wind by the winds at 700 mb--recognizing fully that the choice was entirely arbitrary--but accepting also the complete impracticality of attempting a more rigorous definition of the wind field for routine multiday modeling applications.

Where constraints intrinsic to practical modeling applications are not overiding issues, the use of sophisticated dynamically-constrained wind field models is being vigorously and successfully pursued--for example, the complex mesoscale modeling systems being developed at the Lawrence Livermore Laboratory. Sherman (1978) has described MATHEW, a terrain-consistent variational wind field generation model, and Lange (1978) has used MATHEW output in driving the regional version of the ADPIC model for verification against regional tracer studies. Such wind field synthesis programs represent the leading edge of the state of the art--but require computer resources that fall well beyond those practical or attainable for routine multiday impact assessments. For the present and in the near term, then, we will often have to select, arbitrarily, the mean transport wind levels and the wind field interpolation schemes.

Over long transport distances, a small error made in predicting or estimating initial plume trajectory direction can often mean a completely missed receptor far down wind. Figure 8 illustrates just how long the initial error can grow to be, assuming straight line flow thereafter (hardly a realistic assumption, in general). The figure shows the plume lateral (crosswind) trajectory error as a function of initial directional error and downwind distance, and compares the lateral error to the characteristic half-width of the plume (taken as 2.15 σ_y) as a function of PGT stability class. Under neutral stability the plume spread may be much larger than the error made in estimating initial trajectory, especially if the vertical wind shear is small. Under stable conditions, however, in which the vertical directional shear of the wind may be very large and the plume quite narrow, the model results will be much less forgiving of an initial error in trajectory. Here, as elsewhere, it is important to keep in practical perspective the relative importance of different sources of modeling uncertainty: if in a given application the wind directional information is unreliable, it cannot matter very much if the stability is 'off' by one class, say; nor for that matter, whether we characterize vertical diffusion by a Gaussian or by a Fickian mechanism.

As suggested above, a major source of trajectory error, one that is now beginning to receive attention in long range transport modeling, is the problem of describing the vertical shear (speed and direction) of the mean wind. Sykes and Hatton (1976) and Smith and Hunt (1978) discuss the wind shear effects on horizontal spread. Under stable conditions, when vertical directional shear may be greatest, the width of the plume after transport on the order of 100 km or more is comparable to that expected from large scale synoptic meandering of the plume trajectory over such transport times. Under neutral to unstable conditions (as Smith and Hunt (1978) point out), the effects of wind shear are much smaller, but under strong penetrative convection plume material may be transported upwards to levels at which the mean wind differs significantly from the wind in the mixed layer below. Unfortunately, most trajectory models of the plume and puff variety cannot economically account for vertical wind shear effects (but cf. Sheih 1978a). Where these effects will exert the dominant influence over regulatory decisionmaking, there may be no recourse but the fully three-dimensional vertically-layered numerical grid approach (except that such calculations become quickly cost-prohibitive for other than episode-type analyses).

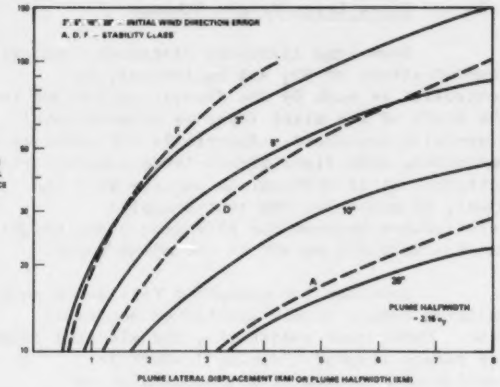


Figure 8 Plume Lateral Trajectory Error versus Plume Half Width

The papers of Draxler (1979), Sykes and Hatton (1976), and Smith and Hunt (1978), among others, illustrate well the profound influence played by the choice of wind field scheme on regional scale predictions. A particularly graphic recent example is offered by Liu and Wojcik (1979). Using their single-layer grid model they compared predicted 3-hour average ground-level concentrations over the Northern Great Plains using (a) subjectively analyzed wind fields prepared by an experienced synoptic meteorologist, and using (b) smoothed 850 mb geostrophic wind fields. They found, in many instances, concentration differences more than twofold. Evidently, the specific wind field scheme chosen for regional and long range flow characterization will have profound (if not always recognized) implications for permitting use of the model; and if the choice of wind field scheme is at least arguably ambiguous even for terrain and flow conditions as well-behaved as those of the Northern Great Plains region, how much more problematic yet is the choice of wind field scheme in truly rugged terrain?

5.2 Mixed Layer Height Schemes

Over long transport distances, ambient concentrations of SO_2 and sulfate may be controlled as much by the diurnal variations in the depth of the mixed layer as by horizontal dispersion processes. Especially for accurately describing near-field ground-level concentration patterns, it is critical to capture with the model, if possible, the ever-changing relationship between the effective plume height and the vertical extent of the mixed layer.

Consider for example a tall stack source emitting into a stably stratified nocturnal flow. Under such conditions, the elevated plume may remain trapped ("ducted") aloft for considerable transport distances with no appreciable ground-level impact; if the model assumes the plume is uniformly mixed, immediately outward from the source, the resultant ground-level pattern will be completely wrong. (The influence of a plume fumigation mechanism within a mesoscale puff model is illustrated in Bass et al. (1979a)). We would have to judge that for many routine regulatory applications, especially to PSD Class I increment consumption, a long range transport model that does not permit the separation of plume vertical spread from that of the mixed layer (or what might be even worse, assumes a climatologically fixed mixed layer height) is inadequate--even for so-called "screening" applications.

Because of the sensitive relationship between plume height and mixed layer height, it is not necessarily the case that increasing the height of the mixed layer will result in a decrease in the ground-level concentrations--certainly not, if an increased mixed layer height causes an elevated plume to be fumigated to the surface. In the far field of a source, plume concentrations are usually greater if the mixing depth is lower, and conversely, because plume elements far from a source will usually have experienced at least one diurnal mixing depth cycle and will have been fully entrained. In contrast, because plume entrainment is episode-specific, no clear relationship can be assumed to exist in the near field of a source.

It is therefore important to consider how the various models differ in their representations of mixing height variations and interpolation schemes. Some models, the EURMAP-1 model for example, use spatially uniform climatological values--but only for long-term averages. Other models, for example the short-term ATAD model, use spatially and temporally varying values. The MESOPUFF model updates the mixing height field hourly, as the larger of a mechanical mixing depth or a convective mixing depth (Benkley and Bass 1979b).

In Heffter's ATAD model, the daytime transport layer depth is calculated as the height in the critical inversion layer where the temperature is two degrees above the temperature at the inversion base. At night, the transport layer depth is $2\sigma_z$, computed from the vertical eddy diffusion coefficient as $2(2\ K_z t)^{1/2}$, using an average nighttime value of K_z.

5.3 Plume Transformation and Removal

The ultimate fate of sulfur oxides in the atmosphere over long transport scales is not very well known. Indeed, major programs such as MAP3S, EPRI/SURE, STATE, OECD LRTAP, and other programs are expected to improve significantly our understanding of plume transformation and removal mechanisms over mesoscale and greater transport distances. It is therefore instructive to look briefly at some of the different approaches taken in the various models to the description of SO_2 and SO_4 transformation and removal.

The ASTRAP model, for example, incorporates dry deposition rates that depend on season and time of day. The SO_2 to SO_4 transformation is linear, but the rate is seasonally and diurnally dependent. By contrast, the PNL regional model assumes constant rates for dry deposition and for SO_2 to sulfate transformation; but the wet removal rate, taken as linear with precipitation rate, is shown to depend critically upon whether precipitation is represented as (a) uniform and constant over the entire computational region; (b) constant but spatially variable; or (c) varying hourly with space and time (Wendell et al. 1977).

The choices of plume removal mechanisms in the model can depend on whether the model is intended for long-term or for short-term impact assessments. For example, the EURMAP-1 long-term model uses a constant dry deposition rate; but its short-term analogue, EURMAP-2, uses dry deposition rates that depend on the height of the puff element above the surface. In many models the dry deposition rate depends on the ground-level concentration, because the surface flux is parameterized in terms of a deposition velocity. Here again, if the ground-level concentrations are misrepresented by the model (because the plume is prematurely mixed to the ground when it shouldn't be, or conversely), the deposition rate will be significantly different, and will therefore charge the subsequent evolution of the plume.

The dry deposition rates may be made to depend on the underlying surface: Lavery et al. (1980), have found, using the advanced EPRI/SURE regional grid model (SURAD) improved model performance using a dry deposition mechanism that varies with the underlying surface vegetative cover and with diurnal variations in turbulent eddy stresses.

Virtually every modeling group has tested the sensitivity of their models to variations in deposition and removal terms - but to our knowledge, the work of Lavery et al., and Johnson et al. (1979) are among the few studies that seek to verify the effects of such model variations with actual field data.

The transformation rate of SO_2 to SO_4 in a plume is, it is becoming increasingly clear, a sensitive, complex function of the air mass characteristics - especially the relative humidity and temperature. Hidy et al. (1976) and more recently Tong (1979) have explored this relationship using the SURE data base with models that assume linear first order conversion rates.

More elaborate, nonlinear SO_2-SO_4 transformation mechanisms have been proposed, based on detailed insitu plume observations--of the Labadie plume, for example. But for present or for near-term regulatory applications a prospective user is not likely to have sufficient information on which to base a source-specific nonlinear plume transformation rate.

Another of the outstanding problem areas in long range transport modeling, and especially in view of the increasing concern over acid precipitation, is that of wet removal. Rainfall is extremely efficient at removing ambient SO_2, and during active precipitation, washout easily dominates removal (Smith and Hunt 1978). Various prescriptions are used for modeling washout in terms of rainfall rate and characteristic scavenging efficiency; these offer at least formally the capability to describe wet removal. But because precipitation rates can be highly irregular, spatially-limited, and episodic, especially during active convective situations, it is difficult if not impossible to categorize rainfall rate on a scale adequate to describe the fate of a plume, especially in its early time history. Several studies have indicated the important episode-like nature of the plume rainout problem, and point thereby to the difficulty of the acid precipitation modeling problem: the significant uncertainties in plume transport and mixing are compounded by the uncertainties in the nature, extent and duration of the precipitation pattern experienced by a plume in the course of long range transport. Because of the disproportionate importance of only a few episodes to long-term acid precipitation burden, it will be important to make modeling studies with spatially and temporally well-resolved information on precipitation rates during characteristic episodes. If wet deposition is to be represented systematically on a regional basis, we will need objective schemes for rainfall rate analyses--schemes that have yet to be developed, at least for practical mesoscale modeling. Here again, we may look to the LAMPS program for significant new developments (but not, realistically, in the near term).

6. MODEL SENSITIVITY AND MODEL VERIFICATION STUDIES

Until quite recently, attempts at verifying trajectory models at long transport distances have been severely hampered for lack of suitable inert tracers detectable at infinitesimal concentration levels, yet easily distinguishable from natural or anthropogenic background concentration levels. Of late, however, the NOAA/ARL Idaho Falls laboratory in particular has pioneered in the use of special tracer techniques for long range transport measurement programs. Techniques have also been developed to exploit a tracer of opportunity, (Krypton-85), emitted routinely from the Savannah River Plant. Detailed descriptions of more recent experiments with SF_6, CD_4, and Krypton-85 releases from the Savannah River Plant, and releases of SF_6, heavy methane, and fluorocarbons from the Idaho National Engineering Laboratory are given by Draxler (1979).

Numerous trajectory modeling studies have been conducted to compare predicted and observed long-term average regional-scale SO_2 and SO_4 concentration budgets - in Scandinavia, over western Europe, in the eastern United States, and elsewhere. The models used - simple multisource trajectory models, or grid models of modest spatial resolution - have met with variable to good success on these scales (see below).

But, in proposing to use regional scale models for regions in which tracer data or other adequate field monitoring concentration data does not exist for model verification, modelers have often been forced to rely solely on model sensitivity studies. Not only for lack of adequate data bases, but also to better understand the capabilities and limitations of their techniques, most modeling groups have made extensive model sensitivity tests. Unfortunately, only a fragment of this important collective experience has been reported.

As study of the long range transport modeling literature will quickly reveal, it is frustrating to model sensitivity results of different modeling groups, because no two models begin from substantially identical physical assumptions, initial conditions, spatial and temporal resolution, meteorological, emissions, and so forth. Different workers have tended to stress, perhaps as a matter of personal predilection or common interest, different areas of long range model sensitivity.

For example, Wendell and colleagues (Wendell et al. 1977) have looked in detail at the effects on their model of time averaged versus real time precipitation in wet removal of SO_2; they found that time averaged precipitation causes significant overremoval of SO_2, and very different spatial distributions of ambient sulfate levels. And Powell et al. (1979) found significant differences between model results under assumptions of variable, or constant, stability class; these were related to differences in plume aging and degree of prior vertical mixing.

Many long range transport studies (e.g., Liu et al. 1977, Bass et al. 1979a) have illustrated that model sensitivity to changes in one parameter cannot be assumed independent of changes in another parameter. For example, Powell et al. (1979) found that changing from constant neutral stability to variable stability caused changes in relative deposition of SO_2 and SO_4 because, in their model, dry deposition is proportional to ground level concentrations, but wet deposition is proportional to the vertical layer averaged concentration field. Shannon has pointed out the importance of including diurnal variations in vertical stability to describe the time dependent coupling of elevated plumes to surface boundary layer processes. Rao et al. (1976), Johnson et al. (1978), Powell et al. (1979), and many others have described the relative sensitivity of their modeling results to different choices of plume transformation and removal, wind speed and mixing height assumptions, grid resolution, and other model parameters and input variables. But, to date, it would seem no group has undertaken fully the task of systematically comparing the performance of the various suggested modeling approaches with a consistent set of experiments beginning with the same meteorology, the same emissions, the same observed ambient air quality, geometry, resolution, etc., and as similar a choice of base case model parameters as is possible. (It is currently being planned, however, to conduct a comparative analysis of the performance of three trajectory type models and one grid model developed at the various National Laboratories against the EPRI/SURE and MAP3S data bases (personal communication, Paul Michael)).

It is not far amiss to describe the model sensitivity testing of regional scale and long range transport models as anarchic--many groups are repeating much the same kind of experiment, but no consistent, uniform protocols exist by which to compare and evaluate the respective performance of different models.

More generally, one of the crucial issues for verification of regional scale models is that of suitable performance evaluation criteria. In recent EPA-sponsored studies, several such criteria have been advanced (Hayes 1979; Hillyer et al. 1979; Gelinas and Vajk 1979). Other groups, in particular the EPRI Plume Model Validation study currently underway, are also confronting squarely the issues of appropriate model performance evaluation criteria. Until, however, a reasonably broad consensus for acceptability criteria emerges among the long range transport modeling community, it will remain difficult to prove convincingly that one or another model is clearly preferable for a given regulatory application. Indeed, as perceived by interested regulatory groups, the choice of long range transport model approaches may appear to be a matter of subjective taste; and arguments about physical realism may not be especially persuasive to those not familiar with the complex nature of the long range transport problem.

Under NOAA sponsorship, the ERT group has made an attempt at systematizing such a set of comparison criteria for long range transport models, especially for model sensitivity studies. They compared the respective sensitivity of the plume segment, puff, and grid (moment-method) modeling approaches under identical meteorological and resolution conditions. These are reported in considerable detail in Bass et al. 1979a.

6.2 Verification of Long-Term Models

It is fair to say, we believe, that the verification history for long range transport models is as yet fragmentary, although efforts have accelerated considerably within the last several years. Some verification is available for long-term average regional scale models. For example, Mancuso et al. (1979) have reported on comparisons of modeled and observed monthly-averaged ground-level concentrations of SO_2 and SO_4 for the western European LRTAP data base using the EURMAP-1 model. They obtained correlation coefficients ranging from 0.7-0.8 for SO_2 and from 0.6-0.7 for SO_4. The regional distribution patterns seem qualitatively reasonable.

Heffter et al. (1979) examined a 2-1/2-year long set of weekly average samples of Krypton-85 measured at 13 locations at distances of 30-150 kilometers from the Savannah River Plant. He used both his long-term model (version A) and his short-term (H) model (forerunners of the ARL-ATAD model). His results show considerable scatter, but no systematic model bias.

Pendergast (1977; 1979) also looked at this data base and compared monthly and 10-hour averages using Kern's segmented plume model. Pendergast's results suggest that hourly stability class rather than constant stability made little difference--but, it will be noted, these are near-surface releases (63 meters).

Meyers et al. (1979a), using essentially a version of Heffter's puff model with more complex vertical diffusion (a forerunner of AIRSOX), looked at monthly average concentrations of SO_2 and SO_4 using the NADB and EPRI/SURE data bases for the eastern United States. They report correlations of 0.7+ for SO_2, and 0.6+ for SO_4, with the best correlations as high as 0.8 for SO_4 during some months.

Shannon (1979a) describes long-term average verification studies with the ASTRAP statistical trajectory approach, with generally favorable results.

And finally, the extensive model verification history compiled in the EPRI/SURE program has demonstrated crucial relationships between characteristic meteorology and episodes of high sulfate concentrations (Hidy et al. 1976; Hidy et al. 1979; Lavery et al. 1978).

Recently, McNaughton (1980) has provided another set of comparisons of the SURE and MAP3S data base observations with long-term predictions of the PNL regional model; his work emphasizes the importance of studying plume transformation and precipitation scavenging regimes.

Recognizing present and near term limitations, it is important nevertheless to assemble as thorough a verification history as possible for long-range transport models that may be prospectively valuable for regulatory and other practical applications. It is likely that the long-term average model approaches will be verified sooner than the short-term models--for all the problems common to verification of conventional short-range models, exacerbated by the special considerations of long-range transport.

6.3 Verification of Short-Term Models

The verification history of available short-term models is skimpy at best, but improving rapidly. Heffter (1977) has reported on the use of his short-term model for verification of a few selected multiday episodes of Krypton-85 transport over scales of 1,000 km. The agreement is encouraging but much more work is needed. The "H" version of his model, with somewhat better physics than the "A" version, found 50% of calculated weekly averages within a factor of 2 of observations; and on the order of 90% within a factor of 10. But, Heffter points out, the improvement can be attributed essentially to use of hourly average rather than monthly average release rates. He emphasizes that verification success is largely a question of the availability of adequately resolved wind data.

Mancuso et al. (1979) have recently reported on some preliminary results of their short-term EURMAP-1 model against the LRTAP data base - for one daily average comparison test. The qualitative comparison appears quite reasonable but hardly conclusive.

The most persuasive verification experiment reported to date for a short-term average mesoscale trajectory model, at transport distances on the order of 100 km, are the recent Draxler (1979) results. Especially noteworthy are the differences observed in model calculations for different choices of wind data. Draxler found that surface winds did poorest; wind obtained hourly from a 60-meter tower, although not local, provided much better estimates of mean transport; and hourly surface winds adjusted to describe a layer average wind field provided the best correspondence between actual time of arrival and duration of the tracer material.

Draxler's conclusion about the required sampling density for wind station data (on the order of 25 km) points to the real difficulty confronting the proposed use of mesoscale transport model under considerably less well-resolved wind field conditions. He does suggest, however, that for cases of relatively constant and uniform flow fields, the wind field resolution available from the surface wind data network (on average, on the order of 100 km) may be adequate.

We may conclude, overall, that the verification of a short-term trajectory model is a major undertaking, one that a regulatory group contemplating the use of the model for a new region will probably not have the resources to undertake. The burden will remain with the principal model development groups. We understand that the SRP data base will be made available to other interested users for model verification purposes (personal communication J. Heffter) and all groups proposing to submit long range transport models for review as candidate Guideline models should be strongly encouraged to avail themselves of this opportunity.

7. PRACTICAL LONG RANGE MODELS FOR REGULATORY APPLICATIONS

The picture painted above may appear bleak; it is not. Much progress, more than was anticipated even a few years ago, is being made to refine and validate long-range transport models, and more will be forthcoming. Yet the present regulatory requirements for near-term decision-making about source impacts at long transport ranges, and the increasing emphasis on secondary effects of major pollution sources (acid precipitation, regional haze, and the like), underscore the practical necessity for selecting long range transport models for use now. The available models, all of them, are imperfect - and their respective authors would be the first to underscore their limitations. Yet considerable thought and ingenuity has gone into the elaboration and implementation of at least some of these models, efforts to make the models more flexible, more responsive to a range of possible uses, more efficient and modest in computational storage requirements, and more applicable directly to regulatory-type questions.

If the models are to be generally useful for simulation of regional-scale plume transport and dispersion under different meteorological regimes, they should respond to mesoscale spatial and temporal variations in wind field, mixing depth, and ambient turbulence levels. The input meteorology fields necessary to drive the models must be easily generated from readily available data if the models are to be useful to a wide user community. The dispersion models must have moderate execution time and computer storage requirements if they are to be practical for multiple-source, multiple-day simulation exercises. And the model design should also allow for easy adaptation to future research needs and regulatory applications - thus, the programs should be highly modular, having components that are easily modified or readily substituted.

Other important features that contribute to the usefulness of models for regulatory and other practical applications include: (a) the ease with which model input data requirements, parameters, and algorithms can be changed to facilitiate use for different applications; (b) the ease with which the models can be used to simulate short-term and long average ambient ground-level concentrations; (c) the availability of preprocessing and postprocessing systems to prepare input and to analyze and display model results; and (d) the freedom to change the model geographic extent and spatial resolution and to modify (replace) the meteorological preprocessing system.

Table 1 lists a number of long range transport models in the public domain. The table stratifies the available models by three classes:

- short-term models that may be appropriate for practical regulatory applications;
- long-term models that may be appropriate for practical regulatory applications; and
- a representative sampling (but far from a complete list) of current research-grade long range transport models under active development at the National Laboratories and elsewhere.

The models identified in the first two classes are all generally available (or will soon be) and documented for external users; some of

TABLE 1
CURRENT LONG RANGE TRANSPORT MODELS (USA)

	Type	Name	Sponsoring Group	Reference
(a) Short Term Models for Regulatory Applications	Puff	Atmospheric Transport and Diffusion Model (ARL-ATAD)	NOAA Air Resources Lab	Heffter 1980
	Puff	EURMAP-2	SRI	Mancuso et al. 1979
	Puff	MESOPUFF	ERT	Bass et al. 1979
	Plume Segment	Source-Transport Receptor Analysis Model (STRAM)	Battelle Pacific North-west Lab	Hales et al 1977
	Plume Segment	MESOPLUME	ERT	Benkley and Bass 1979
	Grid	Northern Great Plains Regional Model	SAI	Liu and Durran 1977
	Grid	MESOGRID	ERT	Morris et al. 1979
(b) Long-Term Models for Regulatory Applications	Statistical Trajectory	Advanced Statistical Trajectory Regional Air Pollution Control Model (ASTRAP)	Argonne National Lab	Shannon 1979
	'Square' Puff	PNL Regional Pollutant Transport Model	Battelle Pacific Northwest Lab	Powell et al. 1979
	Puff	ARL-ATAD	NOAA Air Resources Lab	Heffter 1980
	Puff	EURMAP-1	SRI	Johnson et al. 1978
(c) Representative Research and Development Models	Puff/Vertical Finite Difference	Mesoscale Trajectory and Diffusion Model	NOAA Air Resources Lab	Draxler 1979
	Puff/Vertical Finite Difference	AIRSOX	Brookhaven National Lab	Meyers et al. 1979
	Puff	PNL Regional Model	Battelle Pacific Northwest Lab	McNaughton 1980
	Plume Segment	Segmented Plume Model	Savannah River Lab	Pendergast 1979
	Particle-in-cell	ADPIC-MATHEW	Lawrence Livermore Lab	Lange 1978
	Grid	SULFA3D	EPRI/ERT	Rao et al. 1976
	Grid	SURAD	EPRI/ERT	Lavery et al. 1980
	Grid	Regional Transport Model	Teknekron	Mills and Hirata 1978
	Grid	'Regional Supermodel'	EPA Meteorology Lab	Lamb
	Trajectory	Limited Area Mesoscale Prediction System (LAMPS)	Drexel-NCAR-BNL	Kreitzberg and Leach 1978

the models in the third class are also so documented. As emphasized in the Introduction, we believe the short-term models are more urgently required at this time, but one may anticipate increasing demand for long-term average models as well.

The models suggested as for regulatory applications are all, to our knowledge, well-coded, readily-transportable, easy to use and to modify, and in general appropriate tools to begin an analysis of long range transport and dispersion. But it should be emphasized that skilled judgment is still critical in their use and interpretation. The different models have of course different advantages and disadvantages - the prospective user is strongly urged to consult the specific references in greater detail, and, almost of necessity, to speak to the authors as well. (We might note that the ERT models--MESOPUFF, MESOPLUME, and MESOGRID--have been designed from the outset specifically to be practical for regulatory purposes - and are under consideration for wider distribution as EPA UNAMAP "INDIVIDUAL" models.)

By and large, the short-term puff and plume-segment models are economical for simulating one or a small number of sources, but would not be appropriate for simulating large multisource inventories (e.g., heavily industrialized regions). For such problems, grid models are more practical.

The long-term trajectory models, and particularly the statistical trajectory and 'square puff' approaches, are especially economical. If the loss of near-field resolution is tolerable, they may well offer the most efficient path to long term average computations.

Of course, a short-term model can be used to develop long-term averages by "brute force" sequential calculations. There is considerable precedent for this approach in using conventional short range Gaussian plume models. It is indeed possible to do so for a mesoscale trajectory or grid model, but the costs may be prohibitive, especially for a large number of sources. Figure 9 illustrates just such a 'brute force' approach to developing long-term averages with a short-term puff model (MESOPUFF). In all, 24-hour average concentrations were computed for 106 days during calendar year 1977 for the Four Corners region, and these were weighted to simulate the annual average concentrations of SO_2 and SO_4 in the region. Simulations of this magnitude may cost tens of thousands of dollars; if the objective is only to develop long term (but not maximum short term) average concentrations, the long-term average modeling approaches are more attractive.

Finally, the usual spatial resolution of these suggested models is typically on the order of 50 km (± a factor of 2). By and large, their resolution is adequate for α-mesoscale transport problems, and adequate or marginally adequate on the β-mesoscale.

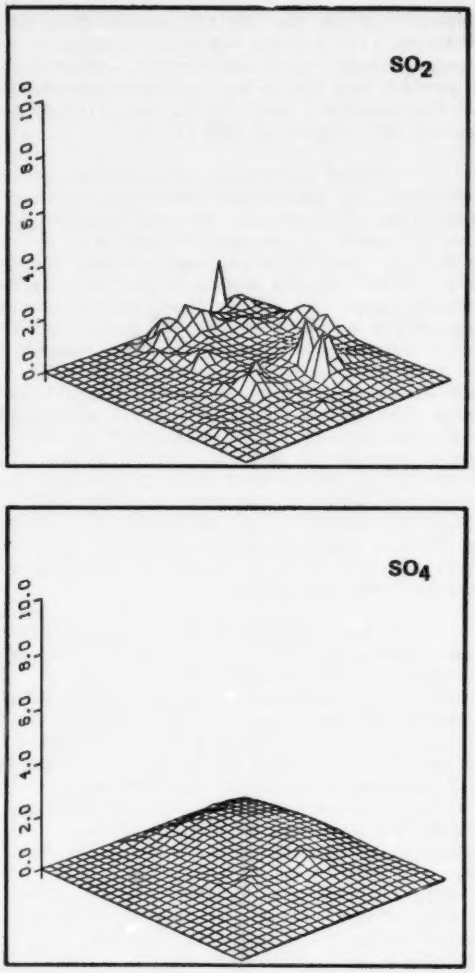


Figure 9 Quasi-Annual Average Ground-Level Concentrations ($\mu g/m^3$)

8. FUTURE DIRECTIONS AND RESEARCH NEEDS FOR LONG RANGE TRANSPORT MODELS

The needs may be put simply:

- standard model verification data sets;
- standard baseline cases;
- standard protocols for model testing; and
- uniform criteria for model performance evaluation and verification.

Meeting these needs, unfortunately, is not simple. The evaluation criteria should, practical experience suggests, reflect intended regulatory usage - that is, the model's ability to predict highest, or second highest local values, not just grid-cell averages or area-wide averages, or total pollutant fluxes and budgets. We alluded earlier to the problem of choosing suitable statistical measures. Conventional measures such as correlation coefficients have inherent problems for regional-scale plume model validation with limited sampling station coverage, given the characteristic spikiness of the primary SO_2 field; but are more successful with the smoother, more regular behavior of the regional and long-range SO_4 fields.

Several workshops have underscored the requirement for additional tracer studies to investigate the chemistry and the physics of plume transport on regional scales, and the need for further studies of regional ambient air quality (not sensibly dominated by local source contributions). To verify short averaging time plume models on the β-mesoscale range, ideally, mobile airborne sampling stations are needed. Even then the sampling problem is formidable. Sheih et al. (1978b) have looked in detail at the mobile network sampling problem; for short-term plume measurements 100 km downwind from the source, made to characterize the crosswind plume spread σ_y, one would need several hundred independent aircraft traverses under identical atmospheric stability and mechanical turbulence conditions. Experimental resources of this scale are virtually without precedent in the long range transport modeling field.

Better techniques, including satellite plume tracking and real-time prognostic mesoscale wind field prediction modeling, could help greatly in the forecasting of short-term plume movements and near-real-time positioning of mobile samplerstTo promote greater capture statistics. But except for short, intensive field programs like those suggested by Sheih et al., for practical reasons an observational network for verification of a mesoscale trajectory model will be fixed. Because the effects of large scale plume meander will probably dominate small scale turbulent diffusion for many flow situations, it may be hard to develop reliable estimates of relative diffusion - so for fixed station networks, the greatest emphasis should probably be placed on verifying crosswind-average or lonr-term average predictions. Fixed station networks may be practical for long-term average trajectory models, or for models that essentially ignore horizontal diffusion, or for models in which relative small scale diffusion and large scale meander are not specifically distinguished.

9. FUTURE DIRECTIONS IN LAGRANGIAN LONG RANGE TRANSPORT MODEL DEVELOPMENT

Because the predictive success of regional and long range trajectory models is ultimately constrained most by the spatial and temporal resolution of the available mesoscale meteorological data, in the future regional scale diffusion models may be driven by the output of prognostic fine-mesh numerical weather prediction models. We may hope to understand better on theoretical grounds the key problem of turbulent dispersion in the so-called 'spectral gap' (between the large, energetic baroclinic scales at which available potential energy and kinetic energy are created and the small, turbulent scales at which energy and enstrophy are dissipated). Perhaps it will be clearer, not many years hence, whether the approximations made for plume growth with distance at long travel times are of general or more limited validity (see Gifford 1976 for a review of the basic data and theoretical issues).

As the cost of computation continues to fall we can look forward to models that make much greater demand on large scale computational resources. Such models may become practical for routine assessment use of techniques that are currently much too computer-intensive to be practical (for example, the ADPIC model). We may anticipate that trajectory modules may be embedded routinely within larger grid models to provide subgrid scale detail--already being attempted in a rudimentary way by several modeling groups.

It will remain important to keep in perspective, however, the real constraints on model improvement that arise from imperfect understanding of the physics and chemistry of long range plume transport. As such knowledge accumulates materially, it will support modeling approaches of increasing sophistication--but it is difficult to imagine that more complex models will necessarily be improvements over simpler models until the available mesoscale meteorological data is greatly expanded--either by a denser network of direct observations, or by skillful interpolation with powerful objective analysis or prognostic methods.

For practical applications of long range transport models, it would be very helpful to see the development of interactive "problem definition" programs - for example, as developed for the LIRAQ modeling system - to assist the nonspecialist in using sophisticated models for routine applications.

We should encourage the use of these models to address regulatory questions couched in probabilistic terms--that is, not just "What is the highest or second highest concentration observed in a given time period? but, perhaps, "What is the expected return period for a given concentration value? Impetus is developing for regulatory decisions based on probabilistic interpretations of familiar short range models; it is surely no less appropriate to use long range transport models in similar probabilistic ways.

Some of the research and development models identified in group (c) of Table 1 may, in time, also become practical for regulatory use--especially as the cost of large computations decreases, and large-core machines are more generally available to user groups. Of course, many models should properly remain research models only, as test beds for new concepts and interpretations of turbulent plume transport and transformation over long ranges.

10. RECAPITULATION

Overall, as we have seen, the available modeling resources for practical regulatory applications of long range plume transport and impact are at an 'adolescent' stage: promising but not yet mature. Although the verification history for these models is incomplete at best, they are still better tools for regulatory decisionmaking on long transport scales than are the conventional short range straight-line Gaussian plume models. Important limitations will continue, in the near term at least, to constrain the accuracy and reliability of such models--but the alternative, to continue to use straight-line techniques that are clearly inadequate to describe mesoscale and long range transport, cannot be acceptable. We would suggest that, in the present state of understanding, the use of these models with reasonably conservative parameter choices could well serve as "screening" tools for regulatory decisionmaking--more detailed analyses, with models specifically verified, if possible, for a given region, might be required at a later date.

The need for more frequent, denser meteorological data sets has been emphasized--as has the difficulty of defining the mean wind field (and other meteorological parameters) in regions of rough terrain or under active weather conditions--and has underscored the the prospective importance, in the longer term, of coupled mesoscale prediction/plume transport systems such as LAMPS.

Sensitivity of the available long range transport models to the meteorological, chemical transformation, and removal parameters (not to speak of the intrinsic limitations of one or another numerical scheme) require much more study. Such further work would be more purposeful, we would suggest, if coordinated within a commonly accepted framework of problem definition, meteorology, and physical processes. The need for commonly accepted criteria for model performance verification cannot be overemphasized, and the importance of choosing evaluation criteria *that relate directly to practical end uses of these models* must be addressed with real urgency.

Working with joint purpose, the developers of long range transport model, and those charged with making regulatory decisions can together advance both the state-of-the-art of long range transport modeling, and the national goals of energy development and industrial growth consistent with environmental protection.

Acknowledgment

We acknowledge with gratitude the kind assistance of Dr. Paul Michael, BNL, in providing a preliminary copy of the MAP3S Modeling Abstracts.

References

Bass, A., C. W. Benkley, J. S. Scire, and C. S. Morris 1979a. *Development of Mesoscale Air Quality Simulation Models. Volume 1. Comparative Sensitivity Studies of Puff, Plume and Grid Models for Long-Distance Dispersion Modeling.* EPA 600/7-79-XX, Environmental Protection Agency, Research Triangle Park, NC, 238 pp.

Bass, A., C. W. Benkley, and J. S. Scire 1979b. *Energy-Related Regional Air Pollution In the Four Corners Area, 1977-1987: Simulation Studies with the MESOPUFF Model.* EPA-600/7-79-XX.

Benkley, C. W. and A. Bass 1979a. *Development of Mesoscale Air Quality Simulation Models. Volume 2. User's Guide to MESOPLUME (Mesoscale Plume Segment) Model.* EPA 600/7-79-XXX, Environmental Protection Agency, Research Triangle Park, NC. 141 pp.

Benkley, C. W. and A. Bass 1979b. *Development of Mesoscale Air Quality Simulation Models. Volume 3. User's Guide to MESOPUFF (Mesoscale Puff) Model.* EPA 600/7-79-XXX, Environmental Protection Agency , Research Triangle Park, NC, 124 pp.

Benkley, C. W. and A. Bass 1979c. *Development of Mesoscale Air Quality Simulation Models. Volume 6. User's Guide to MESOPAC (Mesoscale Meteorology Package).* EPA 600/7-79-XXX, Environmental Protection Agency, Research Triangle Park, NC, 76 pp.

Benkley, C. W. and L. L. Schulman 1979. Estimating Hourly Mixing Depths from Historical Meteorological Data. *J. Appl. Meteor.* *18*:772-780.

Bolin, B. and C. Persson 1975. Regional Dispersion and Deposition of Atmospheric Pollutants with Particular Application to Sulfur Pollution over Western Europe. *Tellus* *27*:281-310.

Clark, T. L. and R. E. Eskridge 1977. *Nondivergent Wind Analysis Algorithm for the St. Louis RAPS Network.* EPA 600/4-77-049. Environmental Protection Agency, Research Triangle Park, NC 63 pp.

Draxler, R. R. 1977. A Mesoscale Transport and Diffusion Model. National Oceanic and Atmospheric Administration Tech. Memo. ERL-ARL-64, Air Resources Laboratories, Silver Spring, MD.

Draxler, R. R. 1979. Modeling the Results of Two Recent Mesoscale Dispersion Experiments. *Atmospheric Environment* *13*:1523-1533.

Egan, B. A., K. S. Rao, and A. Bass 1976. *A Three-Dimensional Advective-Diffusive Model for Long Range Sulfate Transport and Transformation.* Seventh International Technical Meeting on Air Pollution Modeling and its Application, Airlie, VA, September 7-10, 1976, 697-714 p.

Eliassen, A. 1978. The OECD Study of Long Range Transport of Air Pollutants: Long Range Transport Modeling. *Atmospheric Environment* *12*:479-488.

Eliassen, A. and J. Saltbones 1975. Decay and Transformation Rates of Sulphur Dioxide as Estimated from Emission Data, Trajectories, and Measured Air Concentrations. *Atmospheric Environment 9*:425-429.

EPA 1978. *Guideline on Air Quality Models*. EPA-450/2-78-027. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC.

Fisher, B.E.A. 1975. The Long Range Transport of Sulfur Dioxide. *Atmospheric Environment 9*:1063-1070.

Fisher, B.E.A. 1978. The Calculation of Long Term Sulphur Deposition in Europe. *Atmospheric Environment 12*:489-501.

Gelinas, R. J. and J. P. Vajk 1979. *Systematic Sensitivity Analyses of Air Quality Simulation Models*. EPA-600/4-79-035. Environmental Protection Agency.

Gifford, F. A. 1976. Tropospheric Relative Diffusion Observations. *J. Appl. Meteor. 16*:311-313.

Gillani, N. V. 1978. MISTT: Mesoscale Plume Modeling of the Dispersion, Transportation and Ground Removal of SO_2. *Atmospheric Environment 12*:569-588.

Goodin, W. R., G. J. McRae, and J. H. Seinfeld 1979. A Comparison of Interpolation Methods for Sparse Data: Application to Wind and Concentration Fields. *J. Appl. Meteor. 18*:761-771.

Hales, J. M., D. C. Powell and T. D. Fox 1977. *STRAM - An Air Pollution Model Incorporating Non-Linear Chemistry, Variable Trajectories, and Plume Segment Diffusion*. EPA 450/3-77-012. Environmental Protection Agency, Research Triangle Park, NC, 147 pp.

Hayes, S. R. 1979. *Performance Measures and Standards for Air Quality Simulation Models*. Report EF78-93R2. Systems Applications, Inc., San Rafael, CA.

Heffter, J. L. 1980. Air Resources Laboratories Atmospheric Transport and Dispersion Model (ARL-ATAD). National Oceanic and Atmospheric Administration, Tech. Memo. ERL-ARL-81. Air Resources Laboratories, Silver Spring, MD.

Heffter, J. L., G. J. Ferber, and A. D. Taylor 1975. A Regional-Continental Scale Transport, Diffusion and Deposition Model. National Oceanic and Atmospheric Administration Tech. Memo. ERL-ARL-50. Air Resources Laboratories, Silver Spring, MD.

Heffter, J. L. and G. L. Ferber 1977. Development and Verification of the ARL Regional-Continental Transport and Dispersion Model. *Proceedings, Joint Conference on Applications of Air Pollution Meteorology*, November 29-December 2, 1977, Salt Lake City, UT, 400-407 pp.

Heffter, J. L., G. J. Ferber, and K. Telegadas 1979. Verification of the ARL Transport and Dispersion Model at 30-150 km. *Preprint Volume, Fourth Symposium on Turbulence, Diffusion, and Air Pollution*. January 15-19, 1979, Reno, NV, 372-375 pp.

Hidy, G. M. et al. 1976. *Design of the Sulfate Regional Experiment (SURE) Vol. I Supporting Data and Analysis*. Report No. EC-125, Electric Power Research Institute.

Hidy, G. M., P. K. Mueller, T. F. Lavery, and K. K. Warren 1979. Assessment of Regional Air Pollution Over the Eastern United States: Results from the Sulfate Regional Experiment (SURE). *WMO Symposium on the Long Range Transport of Pollutants*. Sofia, Bulgaria, October 1-5, 1979. pp 65-76.

Hillyer, M. J., S. D. Reynolds, and P. M. Roth 1979. *Procedures for Evaluating the Performance of Air Quality Simulation Models*. Report EF79-25R. Systems Applications, Inc., San Rafael, CA.

Husar, R. B., J. P. Lodge, and D. J. Moore (Eds) 1978. Sulfur in the Atmosphere. Proceedings of the Internatinal Symposium, Dubrovnik, Yugoslavia, September 7-14, 1977. *Atmospheric Environment 12*:1-796.

Johnson, W. B., D. E. Wolf, and R. L. Mancuso 1978. Long Term Regional Patterns and Transfrontier Exchanges of Airborne Sulfur Pollution in Europe. *Atmospheric Environment 12*:511-527.

Kreitzberg, C. W. and M. J. Leach 1978. Diagnosis and Prediction of Tropospheric Trajectories and Cleansing. *Proceedings, 85th National Meeting, American Institute of Chemical Engineers*. Philadelphia, PA, June 4-8.

Lange, R. 1978. ADPIC - A Three-Dimensional Transport-Diffusion Model for the Dispersal of Atmospheric Pollutants and its Validation Against Regional Tracer Studies. *J. Appl. Meteor*. *17*:320-329.

Lavery, T. F., J. W. Thrasher, D. H. Godden, A. C. Lloyd, and G. M. Hidy 1978. Regional Transport and Photochemical Model of Atmospheric Sulfates. *Proceedings of the Ninth International Technical Meeting on Air Pollution Modeling and Its Application*, CCMS/NATO, Umweltbundesant, Berlin.

Lavery, T. F., R. L. Baskett, J. W. Thrasher, N. J. Lordi, A. C. Lloyd, and G. M. Hidy 1980. Development and Validation of a Regional Model to Simulate Atmospheric Concentrations of SO_2 and Sulfate. Proceedings of the AMS/APCA Second Joint Conference on Application of Air Pollution Meteorology, American Meteorological Society, Boston, MA.

Liu, C. Y. and W. R. Goodin 1976. An Interactive Algorithm for Objective Wind Field Analysis. Mon. Wea. Rev. 104:784-792.

Liu, M. K. and D. Durran 1977. The Development of a Regional Air Pollution Model and its Application to the Northern Great Plains. EPA-908/1-77-001. Systems Applications, Inc., San Rafael, CA.

Liu, M. K. and M. Wojcik 1979. Further Studies of Regional Diffusion Modeling in the Northern Great Plains. EPA-908/1-78-005. Systems Applications, Inc., San Rafael, CA.

Mancuso, R. L., C. M. Bhumralkar, D. E. Wolf, and W. B. Johnson 1979. The Exchange of Sulfur Pollution Between the Various Countries of Europe Based on the EURMAP Model. Preprint Volume, Third Symposium on Atmospheric Turbulence, Diffusion and Air Quality. Raleigh, NC, p. 345-354.

McNaughton, D. J. 1980. Initial Comparison of SURE/MAP3S Sulfur Oxide Observations with Long-Term Regional Model Predictions. Atmospheric Environment 14:55-63.

Meyers, R. E., R. T. Cederwall, and W. D. Ohmstede 1979. Modeling Regional Atmospheric Transport and Diffusion: Some Environmental Applications. In: Advances in Environmental Science and Engineering. J. R. Pfafflin and E. Ziegler (Eds.). New York: Gordon and Breach.

Meyers, R. E., R. T. Cederwall, J. A. Storch, and L. I. Kleinman 1979a. Modeling Sulfur Oxide Concentrations in the Eastern United States: Model Sensitivity, Verification and Applications. Preprints, Fourth Symposium on Turbulence, Diffusion, and Air Pollution, American Meteorological Society, Reno, NV. January 16-18, 1979. pp. 673-676.

Mills, M. T. and A. A. Hirata 1978. A Multiscale Transport and Dispersion Model for Local and Regional Scale Sulfur Dioxide/Sulfate Concentrations: Formulation and Initial Evaluation. Ninth International NATO/CCMS Technical Meeting on Air Pollution Modeling and its Application. August 28-31, 1978. Toronto, Canada.

Morris, C. S., C. W. Benkley, and A. Bass 1979. Development of Mesoscale Air Quality Simulation Models. Volume 4. User's Guide to MESOGRID (Mesoscale Grid) Model. EPA 600/7-79-XXX. Environmental Protection Agency, Research Triangle Park, NC, 118 p.

Nappo, C. J. 1978. Workshop on Long-Range Trajectory-Puff and Plume Modeling of Continuous Point Source Emissions. National Oceanic and Atmospheric Administration, Tech. Memo ERL-ARL-72, Air Resources Laboratories, Silver Spring, MD.

New Orleans 1980. AMS/APCA Second Joint Conference on Applications of Air Pollution Meteorology. New Orleans, LA, March 24-27, 1980.

Niemann, B. K. A. A. Hirata, and L. F. Smith 1979. Application of a Regional Transport Model to the Simulation of Multiscale Sulfate Episodes over the Eastern United States and Canada. Proceedings, WMO Symposium on the Long Range Transport of Pollutants and its Relation to General Circulation Including Stratospheric/Tropospheric Exchange Processes. Sofia, Bulgaria, 1-5 October 1979. WMO No. 538, Geneva, Switzerland. P 337-347.

Pack, D. H., G. J. Ferber, J. L. Heffter, K. Telegadas, J. K. Angell, W. H. Hoecker, and L. Machta 1978. Meteorology of Long-Range Transport. Atmospheric Environment 12:425-444.

Pendergast, M. M. 1977. A Comparison of Observed Average Concentrations of ^{85}Kr with Calculated Values Observed from a Wind Rose Model and A Time-Dependent Trajectory Model. Proceedings, Joint Conference on Applications of Air Pollution Meteorology. November 29 - December 2, 1977, Salt Lake City, UT. pp. 253-254.

Pendergast, M. M. 1979. Model Evaluation for Travel Distances 30-140 km. Preprint Volume, Fourth Symposium on Turbulence, Diffusion, and Air Pollution. January 15-19, 1979, Reno, NV. pp. 648-651.

Powell, D. C., D. J. McNaughton, L. L. Wendell, and R. L. Drake 1979. A Variable Trajectory Model for Regional Assessments of Air Pollution from Sulfur Compounds. PNL-2734, Battelle, Pacific Northwest Laboratory, Richland, WA.

Raleigh 1976. Preprint Volume, Third Symposium on Atmospheric Turbulence, Diffusion and Air Quality. October 19-22, 1976. Raleigh, NC.

Rao, K. S., J. S. Lague, and B. A. Egan 1976. An Air Trajectory Model for Regional Transport of Atmospheric Sulfates. Preprint Volume, Third Symposium on Atmospheric Turbulence, Diffusion and Air Quality. American Meteorological Society, October 19-22, 1976, Raleigh, NC. pp. 325-331.

Reno 1979. Preprint Volume, Fourth Symposium on Turbulence, Diffusion and Air Pollution, American Meteorological Society, January 15-18, 1979, Reno, NV.

Salt Lake City 1977. *Preprint Volume, Joint Conference on Applications of Air Pollution Meteorology*. American Meteorological Society, November 29-December 2, 1977, Salt Lake City, UT.

Scriven, R. A. and B. E. A. Fisher 1975. The Long Range Transport of Airborne Material and its Removal by Deposition and Washout. *Atmospheric Environment* 9:49-68.

Santa Barbara 1974. *Symposium on Atmospheric Diffusion and Air Pollution*. American Meteorological Society, Santa Barbara, CA, September 9-13, 1974.

Shannon, J. D. 1979a. *The Advanced Statistical Trajectory Regional Air Pollution Model*. ANL/RER-79-1. Argonne National Laboratory, Argonne, IL.

Shannon, J. D. 1979b. A Gaussian Moment-Conservation Diffusion Model. *J. Appl. Meteor.* 18:1406-1414.

Sheih, C. M. 1977. Application for a Statistical Trajectory Model to the Simulation of Sulfur Pollution over Northeastern United States. *Atmospheric Environment* 11:173-178.

Sheih, C. M. 1978. A Puff Pollutant Dispersion Model with Wind Shear and Dynamic Plume Rise. *Atmospheric Environment* 12:1933-1938.

Sheih, C. M., G. D. Hess, and B. B. Hicks 1978. Design of Network Experiments for Regional-Scale Atmospheric Pollutant Transport and Transformation, *Atmospheric Environment* 12: 1745-1753.

Sherman, C. A. 1978. A Mass-Consistent Model for Wind Fields over Complex Terrain. *J. Appl. Meteor*. 17:312-319.

Smith, F. B. and R. D. Hunt 1978. Meteorological Aspects of the Transport of Pollution over Long Distances. *Atmospheric Environment* 12:461-477.

Start, G. E. and L. L. Wendell 1974. Regional Effluent Dispersion Calculations Considering Spatial and Temporal Meteorological Variations. National Oceanic and Atmospheric Administration Tech. Memo. ERL ARL-44, Idaho Falls, ID.

Sykes, R. I. and L. Hatton 1976. Computation of Horizontal Trajectories Based on the Surface Geostrophic Wind. *Atmospheric Environment* 10:925-934.

Tong, E. Y. 1979. *Characterization of Regional Sulfate/Oxidant Episodes in the Eastern United States and Canada*. Paper presented at 72nd Annual Meeting, Air Pollution Control Association, Cincinnati, OH, June 1979.

Wendell, L. L., D. C. Powell, and R. L. Drake 1976. A Regional Scale Model for Computing Deposition and Ground Level Air Concentration of SO_2 and Sulfates from Elevated and Ground Sources. *Preprint Volume, Third Symposium on Atmospheric Turbulence, Diffusion, and Air Quality.* American Meteorological Society, Raleigh, NC. pp 318-326.

WHO 1979. *Symposium on the Long-Range Transport of Pollutants and its Relation to General Circulation including Stratospheric/Tropospheric Exchange Processes.* Sofia, Bulgaria October 1-5, 1979.

Wendell, L. L., D. C. Powell, and D. J. McNaughton 1977. A Multi-Source Comparison of the Effects of Real Time Versus Time Averaged Precipitation Data on SO_2 and Sulfate Particulate Removal in a Regional Assessment Model. *Preprint Volume, Joint Conference on Applications of Air Pollution Meteorology*. American Meteorological Society. Salt Lake City, UT, November 29-December 2, 1977. pp. 372-379.

Appendix F

VALIDATION AND PERFORMANCE CRITERIA
AIR QUALITY MODELS

Norman E. Bowne

TRC - Environmental Consultants, Inc.

Wethersfield, CT 06109

1.0 INTRODUCTION

Validation of air quality models has been the subject of intense concern for 30 years in the United States. One of the first expressions of modeling validation was described by the U.S. Atomic Energy Commission in their publication of Meteorology and Atomic Energy in 1955. Other organizations were interested in the development of new models and validation of them. Project Prairie Grass was conducted. The data are part of our fundamental background for Gaussian models. Several programs were conducted by the Department of Defense in the 1960s, as well as the NAPCA St. Louis Study. In the 1970s, the Environmental Protection Agency became concerned with the validity of air quality models because they were being applied to make decisions on control of existing plants and siting of new plants.

In 1980, the interest in model validation has reached beyond the government agencies and is now the concern not only of regulatory bodies and the scientific community, but also of the industries. All are vitally concerned with the proper use of the models and the validity of the models under a variety of situations. Therefore, in this paper, I would like to discuss my concept of model validation as it applies to some of the current air quality problems.

2.0 REVIEW

The Environmental Protection Agency considered model validation during the preparation of the Air Quality Modeling Guidelines, EPA (1978). In the Air Quality Modeling Guidelines, it is specifically stated that it is desirable to validate those models that are being used in response to regulatory requirements, especially if those models are different from standard EPA models. It is my concern that even the EPA models require additional validation because, in our recent work, we have seen that there are areas where those models do not accurately represent transport and dispersion of pollutants in the atmosphere. Yet the models must be used to perform some of the impact assessments that are required by law, assessments that lead to million-dollar decisions.

The Guidelines call for the following iterative sequence of validation and correction:

1. Compare predicted concentration (c) against observed concentration (ĉ).
2. Determine the cause of discrepancies.
3. Correct and improve the data bases.
4. Modify the model, if necessary, to provide better mathematical representation of physical reality.
5. Document what has been done.

In the validation summary of the Modeling Guidelines Workshop report (Roberts 1977), a subcommittee was concerned with several questions regarding validation. How well could a model be validated? We chose to divide the types of models into essentially four classes. First is the ultimate model; that is, the model that predicts the concentration of material in both time and space with some high level of accuracy. We will leave high level undefined at this moment. The second group of models is one based on physical principles, but has not the ability or the input data to predict accurately in time and space. However, it could be expected that the model would predict a distribution of concentrations for a lengthy period of time that would match the observed distribution. The third level of models is one where only maximum concentration or some other special feature of interest would be predicted but would not be capable, especially of predicting the time and space resolution or perhaps even the frequency distribution. The fourth level of models is one which has no background in physics or even statistics and is in the nature of a stop-gap tool. An example is roll back.

While the people working with the air quality modeling guidelines espoused the idea that validation was required, they did not go on and develop specifically what constitutes an appropriate validation technique, leaving that to the people performing the work. Some have picked up on this problem of validation and specific programs were conducted. For example, the Regional Air Monitoring Study and Regional Air Pollution Study in St. Louis were designed to permit validation of existing models and development of new models for an urban area. The EPA, in addition, sponsored work to evaluate their CRSTER model for four electric generating stations (Mills and Record -

1975; Mills and Stern - 1975). In recent work sponsored by the American Petroleum Institute, also to be reported at this meeting, model performance was conducted using field data collected during tracer studies that have been conducted during the past 25 years. There have been numerous other studies designed for individual sources which have been concerned with model validity. In many cases, these have given us additional insight into the models, but frequently have been rejected by regulatory agencies as providing too little data to be convincing.

New programs are now underway to investigate model validity and to develop new models. The major programs are those of the Department of Energy in the Geysers area program in California, the Environmental Protection Agency's Complex Terrain Program, and the Electric Power Research Institute's Plume Model Validation Program. The purposes of these programs are to establish the reliability, accuracy, and precision of air quality predictions derived from various models and to develop new models.

Recently, Systems Applications, Inc., in San Rafael, California, has prepared a two-volume summary on the subject of model validation (Hayes 1979 and Hillyer et al 1979) for the Environmental Protection Agency. SAI was concerned with two major areas; that is, to develop performance measures and performance standards for judging the validity of models. This is entirely appropriate for a regulatory agency that has to use the models for a wide variety of situations. However, here I prefer to discuss only performance measures and leave specification of performance standards to others.

3.0 VALIDATION PROTOCOL

One of the concepts that I believe should be adapted from the EPRI program is that of developing a validation protocol (set of rules and methods) before the program is actually very far underway. In order to have a rigorous validation, it is necessary to have a systematic approach to the validation process. This is particularly true because there is need for unambiguous definitions of the operational air quality measures which models predict and which are used for regulatory control purposes.

Within the EPRI program, we have developed two types of validation: operational and scientific. Operational validation is designed to determine how well the models predict in an operational sense; that is, if the problem is the prediction of the maximum concentration during any year, how well does the model predict the frequency distribution and hence the maximum concentration during that period of concern? An operational validation protocol has been developed and is being coordinated with EPA. Scientific validation is left to the people concerned with improving the performance of models because this deals with the reliability of individual modules such as plume rise and dispersion coefficients.

Two important facets of plume model validation remain for discussion.

1. The definition of operational air quality measures against which plume models are to be validated; and

2. Analytical and statistical methodologies for validation of model predictions versus observed values of the parameters agreed to under (1) above.

One of the important aspects of model validation is whether we wish to have a model that predicts the extreme values well. I have assumed that we do. For example, a Gaussian plume model, by its very nature, is based on average conditions. Even though we fit dispersion functions to extremes such as very unstable conditions or very stable conditions, the problem remains that these are averaged values. If one looks at scatter plots of dispersion as represented by the standard deviation for a given stability class, one frequently finds that there is a wide range as shown in Figure I (Minott and Case - 1976). This immediately means that we have to concern ourselves with the potential error that is inherent in the scatter for any single calculation. Therefore, do we need a model that will predict episodic conditions, or do we need a model that predicts better at some level such as the 95th percentile? Another problem we have is that air quality standards are currently written for short-term periods as the second highest occurence during a given year. The present evaluation used by the Environmental Protection Agency as setforth in the Air Quality Modeling Guidelines is, if one predicts at an array of points, one uses the "highest, second highest" predicted concentration; that is, you look at the second-highest value at each receptor point for the year, in the array, and determine the highest of these overall. As long as we are required to predict the highest or second highest value during a year, I do not believe that we will find any model that will prove to be very satisfactory, given our imprecise knowledge of the workings of the atmosphere.

4.0 MODEL EVALUATION

4.1 Measures of Performance

We must have measures of performance that tell us how well models predict concentrations even if we only use them to compare models with each other. Some measures of performance should probably include how well the modules within a given model work. We are concerned with the transport of the material and the dispersion of the material. If there is a stack involved, we are concerned with the plume rise, and in some cases we may even be concerned with portions of the plume rise algorithm.

We are particularly interested in the zone of maximum concentration, its distance from the source for elevated sources, and what happens to the plume when it is in contact with the ground. We are concerned with characteristics such as position and value of the maximum ground-level concentration, the change of concentration with distance, and the patterns of ground-level concentrations. These are pollutant characteristics of primary operational importance, against which overall model predictions are to be evaluated. However, the module predictions of transport, horizontal, and vertical dispersion and chemistry are also of interest today, and in the future we have to add phenomena like dry and wet deposition.

Our relatively simple portrayal of plume behavior in current models must now be made more complicated by introducing realistic appraisals of the effects that topography and surface roughness have on the plume. These features produce distortions in the transport and dispersion of the plume, and in the case of large variations in elevation, in the point of touchdown. Successful models must contend with these realities, and the data used to validate the models must resolve the effects that I have just mentioned.

SAI developed a number of model attributes, performance measures and performance standards. These are summarized in Table 1. They include measures of obvious attributes such as the peak concentration prediction, absence of bias, prediction of position and pattern (spatial alignment), lack of gross error and temporal correlation which is important for photochemical modeling. Performance measures suggested are ratios, means, standard deviations and correlation coefficients. Performance standards are suggested but not generally quantified.

The EPRI program has suggested similar measures of performance which are listed in Table 2 for operational validation (Hilst - 1978).

To the extent that models produce quantitative predictions of pollutant concentration distributions, and these distributions can be measured, a direct comparison of predicted values and observed values can be made and a quantitative assessment of model performance obtained. These measures have also been recommended in the validation protocol to the EPA.

Once assembled, the basic data identified in either Table 1 or 2 may be manipulated in a variety of ways. For example, historical data on the maximum concentration predicted and the maximum concentration observed may be treated independently in accumulative frequency distribution or subjected to extreme value analysis. Such comparisons permit validation of the model predictions as they are used in estimating any given percentile of maximum concentration, a crucial part of establishing permissible emissions from sources, and one for which valid statistical tests are required.

Having identified these measures of performance (or measures of errors if you like) for model validation, more general requirements may be considered in formulating a validation protocol. The steps in developing a protocol are:

1. Technical Validation - "Does the model perform as intended and is it scientifically sound?"
2. Operational Validation - Does the model "compute the correct values?"
3. Dynamic Validation - Can the model be extended or adapted to other places?

4.2 Technical Evaluation

Since most models are continually being improved, computer software for models is often inadequately documented and cannot be easily transferred to similar computers at other locations, much less different computers at different locations. Permitting the model builder to operate his model for all users eliminates an independent evaluation of the translation of theory to software. Since computer models are very complex statements, a model builder seldom evaluates all possible options. Therefore, it is important for a model to undergo technical validation to insure the computational correctness exists and that there are no implicit statements in the model. This requires a series of independent audits or tests by different model users to determine if documentation is adequate, if theory corresponds to software, and if the data reduction necessary to support the model development can be reproduced.

A model should be initially tested on problems it was designed to represent. Thus, operational validation should focus on problems that satisfy the assumptions and limitations defined by the model builder. Dynamic validation is actually an examination of the model's ability to address problems for which it was not designed. This latter step is important in a complete validation program.

It is useful to classify models into three categories: 1) Simple mathematical statements that can be effectively evaluated by hand or small computers; 2) Simple algorithms that transform a complex set of input into simple or complex set of output; or 3) Complex algorithms that transform a simpler set of input into a complex set of output. The previous cautions on model validation apply to the third category. The other two are usually transparent and can undergo technical validation and dynamic validation by inspection. The classification of models based on their conceptual framework can also provide a guide for validation. A model based on well accepted theory will require technical validation, but may not require operational or dynamic validation. A model based on empirical formulations may not survive dynamic validation. A model with an incomplete theoretical basis and/or involving empirical basis must receive a rigorous validation since it may be tuned for particular events and perhaps cannot be validated for general use. In all cases, it is extremely important to label the model scope of applicability. I would point out that almost all models presently used are based on an evolving empirical basis: Gaussian models on the bases of the dispersion coefficients as represented by standard deviations; and numerical models on the bases of the development of flow fields and eddy diffusivities from observed values.

5.0 STATISTICAL CONSIDERATIONS

EPRI conducted a statistical workshop of validation procedures in January, 1980. The goals of that workshop were: 1) that the statistics should be concentrated on looking for model attributes, not performance standards; 2) they should delineate appropriate quantitative measures; and 3) recommend methodologies.

Guidelines that were previously developed (Hilst - 1978) and considered during the workshop were that several criteria are essential to model validation:

1. the time series of residuals should appear as white noise;
2. the histogram of the differences between predicted and observed values (the discrepancies of residuals) should form a symmetric distribution;
3. the mean value of the residuals should be zero;
4. unusually large residuals should be rare;
5. when plotted against spatial or other exogenous variables, the residuals should be random in magnitude and size;
6. a co-plot of residuals and predicted values should show no pattern;
7. lack-of-fit tests should show that variance of the residuals compares well with the variance of the measurements process; and
8. the coefficient of variance should not be large; that is, the ratio of the standard deviations of residuals to the mean value should not exceed, say, .20 over most of the field of observations.

5.1 Analysis of Data from Standard Air Quality Networks

Professor J. Stuart Hunter at Princeton recommended that "humble" statistics be used first, because the law specifies that we need to predict the second highest concentration. Hunter suggested that there should be consideration of the possible serial correlation of the data. For example, it is always advisable to use end-to-end averages for 3 hours, 8 hours, or 24 hours to reduce serial correlation. Running averages will not really give you any additional information because you have not increased the number of independent pieces of information, only added to the processing problem and created a large number of samples that are not independent. He noted that one should be concerned about missing data in performing averages. It was recommended that the EPA Air Monitoring Guidelines be followed, which requires two-thirds of the data be present for computing any longer terms average. For example, at least two hours are needed to compute a 3-hour average, and at least 16 hours of data are required to compute a 24-hour average.

Statisticians prefer to deal with residuals. If we let y equal the observed value and $\hat{y}$ the predicted value, then the residual, R, is the difference $y - \hat{y}$. Other measures that should be considered are the absolute value of the residual and the ratio of the absolute value of the residual to the measured concentration at any given time period, and some information on the distribution of the residuals such as the standard deviation of the residuals and the correlation coefficient of residuals. One potential problem with residuals is that it is highly likely that the value of the residual is equal to either the observed or predicted value because the transporting wind direction is different from the one that has been measured and applied as an input variable to the model.

Bencala and Seinfeld (1979) suggested an assessment package for establishing air quality model performance. They stress the use of residuals along with analysis of trends, concentration maximum and frequency distribution. They have developed a computer program to perform the analysis.

Much can be learned by looking at scatter diagrams by stations where one plots the observed versus the predicted value and also plotting the residual versus the predicted and the residual versus the observed value. Hillyer and Roth (1980) have designated two types of residual plots (Figure 2). Type 1 plots are used in those cases where the residuals have a one-to-one correspondence with the values of the exogenous variable; that is, for each residual there is a unique value for the variable. An example of this type is a plot of residuals against observed or predicted concentrations. Type II plots are obtained in the case where only one value of the exogenous variable is taken for an experiment (e.g., mean wind speed), or the residuals are being plotted as a function of some broad classification (e.g., experiment number). In the Type II plots, many residual values are available corresponding to a given class and they can be represented as frequency distributions with the mean value indicated. Such a plot permits a visual assessment of both the mean residual which is a measure of bias and the form of the residual distribution as a function of the exogenous variable which illustrates the scatter.

It should be pointed out that in order to apply real statistical tests, it is necessary to know the variance of the observations; that is, if you have two stations side-by-side, what is the variability that is inherent in the observations?

A most useful method for comparing lengthy periods of record of observed and predicted concentrations, especially for Gaussian models, is the cumulative distribution function (CDF). The appropriate statistical test is the Kolmogorov-Smirnov non-parametric test. It should be noted that in order to make a valid comparison the CDF's must be parallel. One way to assist in making the values parallel is to eliminate the zeroes from the data set. Since there may be background concentrations from smaller sources in the measured data, one should eliminate the zeroes by looking at the predicted values, taking out all of those hours, and comparing only the cumulative distribution functions when the model was predicting a value present at the station. Otherwise, one is not likely to have parallel distribution functions and will draw erroneous conclusions. If the model is capable of handling background concentrations, and they can be specified, then it would not be necessary to eliminate the zeroes.

Histograms of predicted and observed values should be plotted as a method of examining the distribution function and one can then see the type of distribution that has to be dealt with.

It was noted previously that one hopes the residuals are white noise. Therefore, a histogram of the residuals should be plotted and they should be tested for symmetry; that is, one would hope to find as many negative as positive residual values. During this plotting exercise, one should pay particular attention to the large residual values and attempt to determine what caused them.

5.2 Analysis of Tracer Data

When dealing with tracer data, there is a more dense network of observations and, hence, one can predict at a larger number of locations. In most tracer studies, one usually deals with arcs of samplers at a variety of distances from the source. If we define the observed concentration on a given arc as $y\ (R, O)$ and the predicted value is $\hat{y}\ (R, \theta)$, then we can compare a number of items that are important to the analysis or measure of validity of the model. One of the first things that should be compared is the maximum concentration at each distance without respect to its angular position on the arc. A second measure is the comparison of maximum concentration without respect to either angular position or radial distance from the source. These permit us to evaluate two types of error. The rotation error which is the angular separation between the predicted and observed maximum concentration, and the translation error which is the arc distance difference between the maximum observed and predicted concentrations. Another worthwhile comparison statistic is the cross-wind integrated concentration.

It has been suggested that an analytical form be fit to the observed data to develop a surface so that one could use that surface as a representation (y') of the observed concentration from which one can then determine the maximum concentration, its angular position, and the distance from the source. It is recognized that there may be considerable difficulty in performing a surface fit because of the large gradients that are frequently obtained at the ground from single tracer plumes. A surface fit introduces an additional error measure in describing the error of the model, because there will be an error caused by the fitting process as well as the error in the measurement of the concentration for comparison to the model.

6.0 RECOMMENDED PROTOCOL

Finally, we recommend a protocol for validation of models. Two separate methods are recommended; one for routine air quality data collected continuously, and the second for tracer tests where large numbers of samplers are used for short periods of time.

6.1 Continuous Data Protocol

At present, the primary operational and regulatory uses of models are directed toward the prediction of:

1. maximum ground-level concentration for various averaging times; and
2. maximum incremental increases or decreases in ground-level concentrations as a function of position from the source for various averaging times.

In many instances, such maximum concentrations or incremental changes are the result of contributions from multiple sources, and the conditions leading to the maximum total concentrations are not those associated with the maximum impact of any one source. Consequently, while it is of great importance that a model predict maximum concentrations from individual sources reliably, model performance evaluation must also assess a model's ability to predict the magnitude and spatial distribution of pollutant concentrations from a given source under a wide variety of meteorological conditions.

One of the goals of most model validation programs is to identify conditions for which improved models are required. Frequently the data used in validation are used for subsequent model development. Since data used in the validation of existing models and the development of new models cannot legitimately be re-used for validation of new models, careful segregation of data sets is necessary in order to reserve data for future validation efforts. While some hold that any data from the same location are not independent, I believe it is useful. For example, Hilst and Bowne (1971) reported 78% predictions within a factor of two at Fort Wayne fitting dispersion coefficients. The data set aside for independent tests showed 61% within a factor of two. Data selection procedures to save an independent sample are discussed below.

From the point of view of protocol, the concept of an "independent" data set is more difficult to treat than two related problems. These problems are:

1. Use of the same validation set for iterative "model tuning." (Such a process can, ultimately, match all observed data to within any desired degree of accuracy if enough model parameters are adjusted.) Once the "model tuning" process has been initiated, the fit of predicted to observed concentrations is no longer an objective test of model performance. The models to be evaluated must be selected, documented, and "frozen" prior to undertaking the validation exercise.
2. Simultaneous statistical inference (especially important for comparison of performance of Model A vs Model B). If a large number of models are tested against a finite validation data set, some models will do better than the "typical" model, some worse (assuming the models differ significantly) - as a result of random "noise" effects.

Analyses with station data should be structured to examine model performance in predicting concentrations at fixed points in time. Averaging times to be considered include the basic one-hour average time unit for model prediction, plus 3-hour and 24-hour periods for which current sulfur dioxide air quality regulations are specified. Initial screening should provide 24-hour time blocks in order to provide 24-hour average data. At least 8 hours of the period should have winds directed toward the station. After the initial screening, data should be selected according to the following criteria:

1. The greatest possible range of dispersion conditions are well represented.
2. Periods for which the measured pollutant emission rate from the plant is available are preferred.
3. Consecutive time blocks or more than 24 hours are preferred.
4. Background (upwind) pollutant concentrations are "well behaved."
5. All required model input data are available.

Data that are selected should be clearly identified and not used again for the same model.

The following tests are recommended for station data:

1. Compare magnitude of maximum predicted and observed concentrations.
2. Comparison of cumulative distribution functions of observed (c) and predicted ($\hat{c}$) concentrations. (Screening the data in the manner described above will effectively eliminate the zero concentration values discussed in the statistics section.)
3. Histograms of predicted and observed concentrations and residuals
   $(r = c - \hat{c})$.
4. If a network of stations are present, provide pattern measures and comparison of positions of maximum concentrations.
5. Plot and compare by single station and for the total network.
   a. c vs $\hat{c}$
   b. r vs c
   c. r vs $\hat{c}$
6. For residuals compute or examine:
   a. average value
   b. histogram and cumulative distribution function
   c. standard deviation
   d. average absolute value
   e. time spent dependent auto correlation
   f. spatial correlation.

Table 3 presents a summary of performance attributes, measures, type of test and remarks based on the recommended tests.

At this time, I have no standards of performance to suggest, but if the same statistical measures are applied to a group of models with the same measured data, the comparison of the models may be performed on a relative basis.

6.2 Protocol for Tracer Data

Tracer data supply a more dense spatial pattern with significantly reduced total hours of data. Many of the tests are different; for example, CDF is limited in hours of data available to generate the plot, but we can expect better definition of maximum concentration and position of maximum concentration.

Data selection should be established on the basis of discrete one-hour time units. As with station data, initial screening should establish that the plume was over the sampling network and that the plume centerline is on the grid. Data should be selected from the screened set to satisfy the following criteria:

1. The greatest possible range of dispersion conditions are represented, assuming the experimental design and validation objectives permit.
2. Consecutive time blocks of three or more hours are preferred.
3. Periods with simultaneous measures of pollutant emission and concentration with tracer concentration are preferred if the experimental design permits.

The tests to be applied to the tracer data includes the following:

1. Comparison of the magnitude of the maximum predicted and increased concentrations.
   a. without respect to position;
   b. by distance from the source.
2. Comparison of location of maximum predicted and maximum measured concentration.
   a. on total grid;
   b. angular position on each arc;
   c. radial position.
3. Pattern measures.
4. Frequency distributions.
   a. y
   b. $\hat{y}$
   c. r

A summary for these tests is contained in Table 4.

Additional tests depend on the data and its adaptability to surface fitting techniques. If a surace can be fit, let c' represent the observations at the same points as c is measured and $\hat{c}$ is predicted. If a surface cannot be fit use the observed value y. The purpose of a surface fit is to provide better definition of the position and magnitude of the observed maximum concentration.

Define the variable z by

$$z_i = \frac{c'_i - \hat{c}}{S_{c'_i}}$$

where $S_{c'}$ is the standard error of the surface fit (or the observed data). Standard statistics can be applied to the new variable z such as finding $\bar{z}$, S_z, $|\bar{z}|$, the CDF for z, its histogram, etc. It should be noted, however, that a large value of $S_{c'}$ indicates a poor fit to the data and the z variable may be of little use in comparing models.

The same statistics as described for the station data may be applied to the residuals. Residuals may be computed in two ways:

1. direct difference of observed and predicted; and
2. difference of fitted data and predicted.

In addition, residuals may be computed after translation and rotation or both and examined as a means for comparing ground level concentration patterns. Again, the same comparisons and plots as described for station data are appropriate.

7.0 CONCLUSION

A systematic validation protocol has been proosed based on past efforts, current regulatory needs for models, and consideration of the type of data likely to be available. The suggested performance measures indicate the ability of a model to predict maximum concentrations, concentration patterns, positions of maximum concentration. In addition, the bias of a model must be measured. The types of tests recommended are comparison of cumulative distribution functions, tests of location errors and tests of residuals.

REFERENCES

Bencala, K.E. and J.H. Seinfeld, 1979: An Air Quality Performance Assessment Package, *Atmospheric Environment*, *13*, 1181-1185.

Hayes, S.R., 1979: *Performance Measures and Standards for Air Quality Simulation Models*, Final Report, Contract No. 68-02-2593. Systems Applications, Inc., San Rafael, California.

Hillyer, R.J., S.P. Reynolds and P.M. Roth, 1979: *Procedures for Evaluating the Performance of Air Quality Simulation Models*, Final Report, Contract No. 68-02-2593, Systems Applications, Inc., San Rafael, California.

Hillyer, R.J. and P.M. Roth, 1980: Personal Communication.

Hilst, G.R., 1978: *Plume Model Validation*, Workshop Summary Report, EA-917-54, Electric Power Research Institute, Palo Alto, California 94304.

Hilst, G.R. and N.E. Bowne, 1971: Diffusion of Aerosols Released Upwind of an Urban Complex, *Environmental Science and Technology*, *5*, 327-333.

Mills, M.T. and F.A. Record, 1975: *Compliance Analysis of Time-Concentration Relationships and Violation of a Single-Source Dispersion Model*. EPA-450/3-75-083. Environmental Protection Agency, Research Triangle Park, North Carolina.

Mills, M.T. and R.W. Stern, 1975: *Model Validation and Time Concentration Analysis of Three Power Plants*. Environmental Protection Agency, Research Triangle Park, North Carolina.

Minott, D.H. and C.R. Case, 1976: *Analysis of Field Experiment Dispersion Data to Evaluate the Split-Sigma versus the Delta-T Method of Estimating Horizontal and Vertical Diffusion Rates*, TRC 71469-02 Final Report for Northeast Utilities, TRC-The Research Corporatin of New England, Wethersfield, Connecticut.

Roberts, J.J., 1977: *Report to U.S. EPA of the Specialists Conference on the EPA Modeling Guideline*, Energy and Environmental Systems Division, Argonne National Laboratory, Chicago, Illinois.

U.S. EPA., 1978: *Guidelines on Air Quality Models*, EPA-450/2-78-027, OAQPS No. 1.2-080, Environmental Protection Agency, Research Triangle Park, North Carolina.

TABLE 1

MODEL PERFORMANCE MMEASURES AND STANDARDS

Performance Attribute	Performance Measure	Performance Standard
Accuracy of the peak prediction	Ratio of the predicted station peak to the measured station peak C_{p_p}/C_{p_m}	Limitation on uncertainty in aggregate health impact and pollution abatement costs†
	Difference in timing of occurence of station peak* Δt_p	Model must reproduce reasonably well the phasing of the peak -- say, ± [illegible] hour
Absence of systematic bias	Average value and standard deviation of the mean deviation about the perfect correlation line, normalized by the average of the predicted and observed concentrations, calculated for all stations during those hours when either the predicted or the observed values exceed the NAAQS $(\bar{\mu}, \sigma_{\bar{\mu}})_{\mathrm{OVERALL}}$	No or very little systematic bias at concentrations (predictions or observations) at or above the NAAQS; the bias should not be worse than the bias resulting from monitoring instrumentation error (-8 percent is a representative value)
Lack of gross error	Average value and standard deviation of the absolute mean deviation about the perfect correlation line, normalized by the average of the predicted and observed concentrations, calculated for all stations during those hours when either the predicted or the observed values exceed the NAAQS $(\lvert\bar{\mu}\rvert, \sigma_{\lvert\bar{\mu}\rvert})_{\mathrm{OVERALL}}$	For concentrations at or above the NAAQS, the error (as measured by the overall values of $\lvert\bar{\mu}\rvert$ and $\sigma_{\lvert\bar{\mu}\rvert}$) should not be worse than the error resulting from monitoring instrumentation error
Temporal correlation	Temporal correlation coefficients at each monitoring station for the entire modeling period and an overall coefficient averaged for all stations $r_{t_i}, r_{t_{\mathrm{OVERALL}}}$ for $1 \leq i \leq M$ monitoring stations	At a 95 percent confidence level, the temporal profile of predicted and observed concentrations should appear to be in phase (in the absence of better information, a confidence interval may be converted into a minimum allowable correlation coefficient byusing an appropriate t-statistic)
Spatial alignment	Spatial correlation coefficients calculated for each modeling hour considering all monitoring stations, as well as an overall coefficient average for the entire day $r_{x_j}, r_{x_{\mathrm{OVERALL}}}$ for $1 \leq j \leq M$ modeling hours	At a 95 percent confidence level, the spatial distribution of predicted and observed concentrations should appear to be correlated

---

*These measures are appropriate when the chosen model is used to consider questions involving pollutants subject to short-term standards. They are most important when the pollutant is also reactive.

†These may not be appropriate for all regulated pollutants in all applications. When they are not, standards derived based on pragmatic/historic experience should be employed.

TABLE 2

Plume or Pollutant Parameter	Measure of Performance
Maximum ground level concentration, χ_{max}	$(\chi_{max})_p - (\chi_{max})_o$ (scalar difference)
Position of maximum ground level concentration, $(r, \theta)_{max}$ (polar coordinates)	$([r, \theta]_{max})_p - ([r, \theta]_{max})_o$ (vector difference)
Maximum ground level concentration at distance r	$(x_{max}(r))_p - (\chi_{max}(r))_o$ (scalar difference)
Position of maximum ground level contration at distance r	$(\theta(\chi_{max}(r)))_p - (\theta(\chi_{max}(r)))_o$ (angular difference)
Pattern of ground level concentration	1. Simple correlation of $(\chi(r, \theta))_p$ vs $(\chi(r, \theta))_o$ 2. $(\bar{\chi}(r))_p$ vs $(\bar{\chi}(r))_o$, $(\sigma_\chi^2(r))_p$ vs $(\sigma_\chi^2(r))_o$, and higher order moments

TABLE 3

PERFORMANCE MEASURES FOR ROUTINE (CONTINUOUS) STATIONS

Performance Attribute	Performance Measure (basis for comparison)	Type of Test	Discussion
Accuracy of the highest predictions (single station)	Compare CDF of observed and predicted concentrations for top 5-10 percent of values.	Test for significant difference in number of events above a stated value. Test for difference of means.	Comparisons based on small numbers of events will be subject to large uncertainty.
	Compare n ($\sim$10-20) highest predicted, observed values.	Test for difference of means.	
	Residuals: as a function of predicted and observed concentrations.	Test for bias - average of residuals $\neq 0$ at high values.	Ineffective test if high predicted, observed values are poorly correlated.
Accuracy of model predictions - overall (single station)	Compare CDF of observed and predicted - all values.	Non-parametric test of similarity of CDF's.	
	Distribution of residuals.	Test for bias(average $\neq 0$). Test for symmetry $\mathrm{CDF}(c-\hat{c})-\mathrm{CDF}(\hat{c}-c)$. Test for normality.	
Model Prediction vs. meteorology	Compare CDF of observed and predicted values for category	Non-parametric test of similarity	Categories must be broadly defined in order to achieve data sets of adequate size.
	CDF of residuals for category.	Test for bias. Test for symmetry Test for normality.	
Comparison of performance between stations	Compare CDF of observed, predicted station-to-station (all values)	Non-parametric tests of similarity	Combine results for "like stations".
	Compare CDF of observed, predicted by meteorology category.	Non-parametric tests of similarity.	

TABLE 4

PERFORMANCE MEASURES FOR TRACER NETWORK RESULTS

Characteristic of Concentration Pattern	Basis for Comparing Predicted and Obs. Values	Test/Comparison	Discussion
1. Magnitude of maximum (highest concentration)	1a. Highest values on tracer grid.	1a. Distribution of residuals Test for bias Test for symmetry Test for normality Residuals vs. obs. or predicted value	Highest values on grid should provide a reasonable estimate of max. concentration when footprint is "contained" by the grid.
		Residuals vs. met. variables Compare CDF of predicted, observed values (esp. at high end)	Same analysis possible for subsets separated by meteorology.
	1b. Highest predicted value at any location; highest observed value estimated via interpolation.	1b. Comparisons analogous to those listed in 1a.	Problem with varying confidence limits on interpolated estimates. Standardized residuals required.
2. Location of maximum (highest) concentration. (R_{max}, θ_{max})	2a. Location of highest values on the grid.	2a. Analysis of residuals (see 1a)	Resolution limited to grid spacings; discrete, not continuous variables.
		Compare CDF of pred., observe values.	CDF analyses only meaningful for radial component.
	2b. Locations corresponding to 1b.	2b. Comparisons as in 2a.	Problem with varying confidence limits.
3. Magnitude of highest concentration at distance R	Highest values on a given arc of the tracer grid.	Analysis of residuals	Crosswind resolution expected to provide adequate resolution. No interpolation warranted.
		Compare CDF of predicted, observed values.	
4. Location of highest concentration at distance R $\theta(R)$	Location of highest observed and predicted values.	Analysis of residuals	Resolution limited to grid spacings.

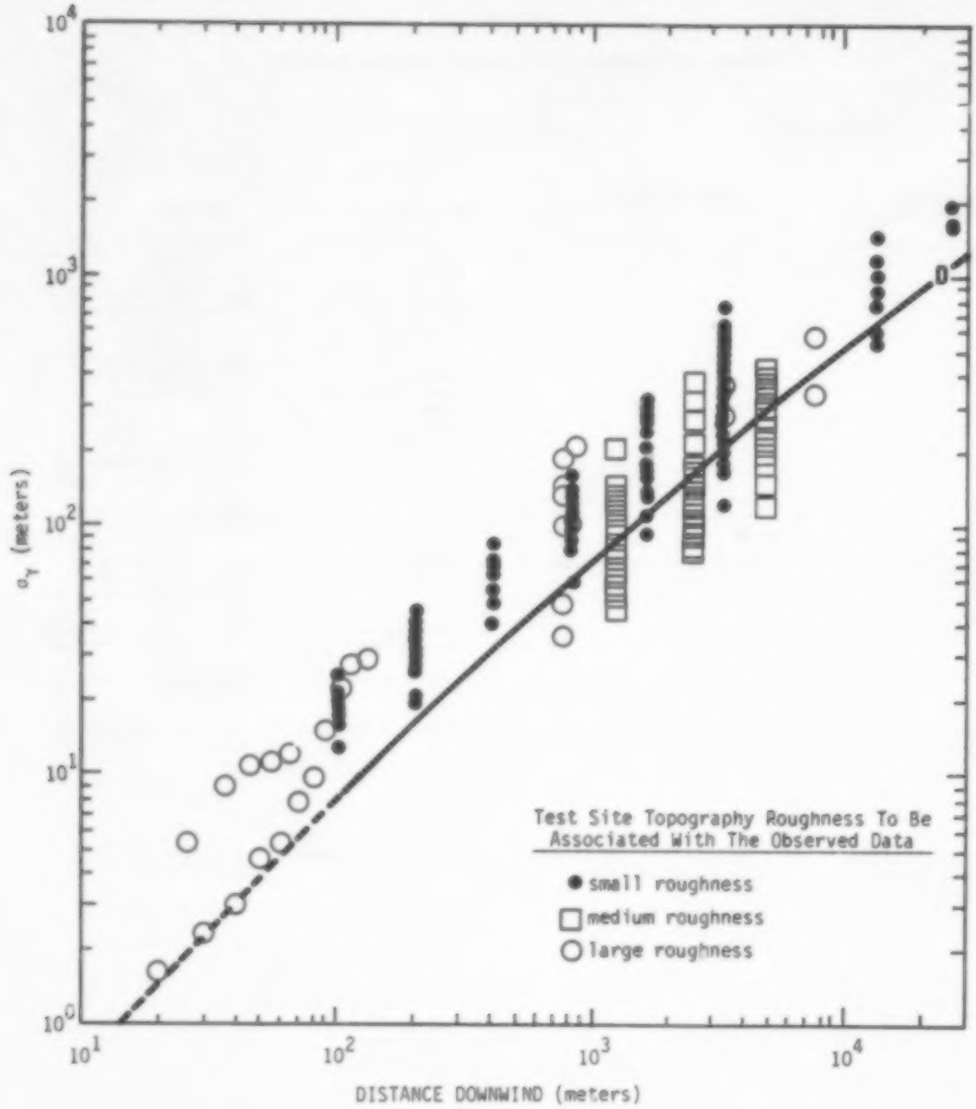
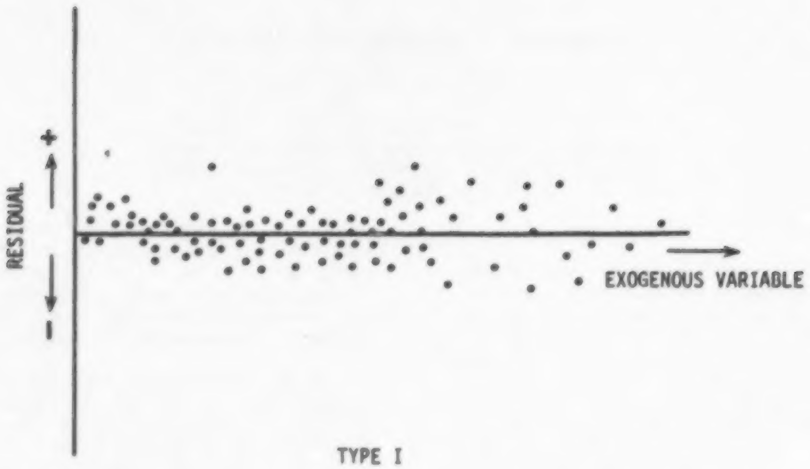
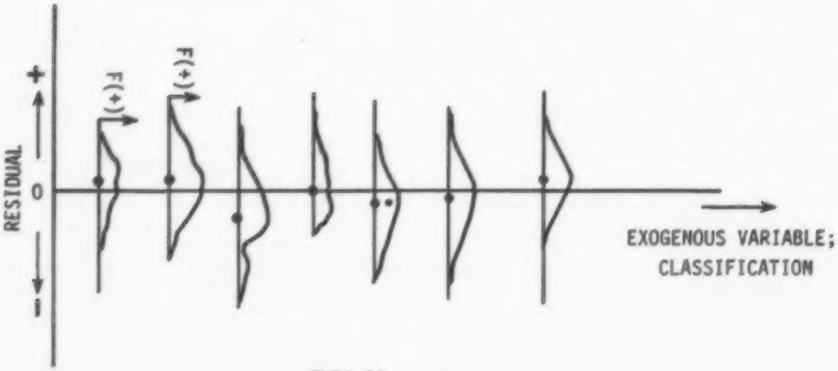


FIGURE 1: OBSERVED SIGMA-Y VALUES & THE PASQUILL-GIFFORD CURVE
FOR STABILITY CLASS D



TYPE I



TYPE II

FIGURE 2 RESIDUAL PLOTS

Appendix G

DISPERSION IN THE VICINITY OF BUILDINGS

R. P. Hosker, Jr.

Air Resources
Atmospheric Turbulence and Diffusion Laboratory
National Oceanic and Atmospheric Administration
Oak Ridge, Tennessee 37830

1. INTRODUCTION

The behavior of effluent plumes near buildings is a concern common to architects, regulatory agencies, and air pollution meteorologists. Reliable answers to the questions of whether a pollution problem exists and, if so, how frequently it may occur, are difficult to obtain. There is relatively little guidance currently available in an easy to apply "handbook" format; instead, reference must usually be made to the technical literature. Thus, for example, the EPA's *Guideline on Air Quality Models* (U. S. EPA, 1978) simply refers the reader to several articles on "aerodynamic downwash"; no explicit modeling recommendations are made. Turner (1969) and the ASME (1973) treat only the case where material emitted from a building vent is fully entrained into its wake. An exception is the excellent article by Meroney (1979), which suggests calculation methods for a variety of cases often encountered in practice.

The present paper is a sweeping condensation of a review article on flow and dispersion near obstacles (Hosker, 1980). A "Handbook on Atmospheric Diffusion" (Hanna *et al.*, 1980) covering these topics, as well as a wide range of other phenomena important to dispersion calculations, is being prepared. Detailed discussions of phenomena described below may be found in these references.

The present text begins with a summary of the flow field characteristic of an isolated block-like building placed perpendicular to the wind. To permit comparison between the discussions of flow and concentration patterns, the material on effluent dispersion has been arbitrarily divided into situations where the pollutant source is upwind (by a few building heights, say) of an obstacle, close to or on the obstacle, or downwind of the cavity behind the obstacle. Surprisingly little information seems to be available on the last topic, despite its obvious importance. Since buildings are notorious for appearing in clusters rather than singly, a brief description of flow and concentration patterns near groups of structures is included. Recommendations for additional contributions and research are given in the text wherever appropriate.

2. SUMMARY OF FLOW NEAR BUILDINGS

An understanding of flow patterns near buildings and the reasons for their existence is central to comprehending and predicting the resulting effluent concentration fields. A recent model of such flow is sketched in Figure 1; the pattern is considerably more complex than was recognized previously (e.g., ASME, 1973). The major phenomena may be summarized as follows.

A deep atmospheric boundary layer is normally incident on a block-like structure. The velocity gradient in the approach flow induces a downward directed flow on the front face; at the ground, this flow moves away from the building in the upwind direction, causing the approach flow to separate from the ground surface some distance upwind. The result is a standing eddy in front of the building. The size and strength of this vortex depend upon the curvature of the wind speed profile relative to the building height, and on the width (W) to height (H) ratio of the structure. For very wide obstacles, wind tunnel results suggest that the main flow contacts the face of the obstacle above the eddy somewhere between 0.5 H and 0.8 H, with 0.6 H being the most commonly observed value. For W/H near unity, laboratory work suggests the same range, depending on the incident wind profile; however, a value near 0.75 H seems the most likely (e.g., Corke and Nagib, 1976). For obstacles much taller than they are wide, Baines (1963) found the frontal eddy reached to a height of about W. The point of flow separation from the ground upwind of a three-dimensional body has not yet been universally specified. Frost and Shahabi (1977) found in field measurements that, at least for a fairly wide building, the separation line was roughly 0.9 H in front of the obstacle. Wilson and Netterville (1978) observed perturbations in incident concentration profiles to first appear about 1.3 H upwind of a model building.

Since the flow separation and reattachment points on the ground and building face are fixed by aerodynamic force balances rather than by geometry, their locations will fluctuate with perturbations in the overall flow field (Hunt, 1975). Such perturbations can be introduced both by ambient turbulence or unsteadiness in the approach flow, and by the feedback effect of turbulence generated in the wake of the obstacle (Norman, 1972). Hence the upwind vortex can be anticipated to be rather unsteady. This vortex interacts with the incident flow to wrap about the building and trail off downwind, as indicated in Fig. 1; fluctuations in its strength and location will therefore lead to perturbations of the wake flow, with subsequent feedback to the upwind vortex. Prediction of the location and characteristics of this vortex-wake interaction is likely to be possible only in a time-averaged sense. It is interesting that laboratory work has shown that the frontal/trailing vortex may actually be a system of many alternating-rotation vortices (Fig. 1 indicates three); in field work, only one seems to have been observed thus far.

Along the upwind edges of the building, the flow is unable to turn sharply enough to follow the roof and sides, and consequently separates from the obstacle. Air immediately adjacent to the roof and sides moves in the upwind direction in these separated zones, and a recirculating flow is formed. The boundary between it and the outer flow is a turbulent shear layer which curves inward toward the body; if the obstacle is sufficiently long, these layers contact its surfaces, and the flow reattaches to the roof and sides. When this occurs, the flow downwind of the reattachment line will be directed toward the rear of the structure. An examination of the literature suggests that reattachment will usually occur in atmospheric flows if the length to height ratio L/H is unity or larger, and the building is not too wide (Hosker, 1979). Reattachment may be delayed or may not occur if W/H becomes large. Wilson (1979) has examined the roof recirculation zone in some detail. Let ζ_b be the smaller of H and W and let ξ_b be the larger. By defining the characteristic length scale for a building as $R \equiv \zeta_b^{2/3} \xi_b^{1/3}$, Wilson finds the roof cavity extends a distance $L \cong 0.9\ R$ from the upwind roof edge, and reaches a maximum height of about 0.22 R above the roof at a distance R/2 from the upwind edge (Fig. 2).

Immediately behind the structure the situation becomes very complex, and earlier pictures of a relatively simple recirculating "cavity" flow bounded by a turbulent shear layer appear to be inadequate. The horizontally-oriented frontal eddy is forced by the incident wind to wrap around the sides of the object and trail off near ground level as a counter rotating "horseshoe" vortex pair. In the lee of the obstacle, near its lateral edges, are a pair of vertically-oriented vortices; these interact with the main flow field near roof level, and also trail off downwind, in this case as an elevated counter rotating pair. The very complex flow field of Fig. 1 has emerged in recent years from the work of Morkovin (1972) and his colleagues, Britter *et al*. (1976), Woo *et al*. (1977), and Hunt *et al*. (1978). Two significant features of this model should be noted: (1) the streamlines "bounding" the cavity zone are *not* the separation streamlines, and (2) transfer of material into and out of the cavity is not limited to turbulent transport across this bounding interface. Instead, material can be *advected* into the recirculation zone along the *entering* streamlines, and can be removed via the rising spiral vortices generated near the lateral lee edges of the obstacle.

Information on the behavior of the horseshoe vortex even close to the building is rather sparse. Ogawa (1973) and Frost and Shahabi (1977) performed a limited amount of flow visualization work during their field experiments; their results only verified the presence of organized vorticity near the edges of obstacles, but did not provide data on their paths, velocities, or persistence. Colmer's field measurements were also suggestive of such vortex structure in the wake. There was some indication in Colmer's (1970) work that the wake effectively rose off the ground as it moved downwind. Similar behavior has been reported in wind tunnel tests (Lemberg, 1973; Hansen and Cermak, 1975). Penwarden and Wise (1975) studied the line of maximum wind speed (y_{ms}) due to the horseshoe vortex wrapping around a rather thin (small L) model building; relative to the wake centerline, their data agree fairly well with the relation $y_{ms}/W \cong 0.5 + c\sqrt{x/W}$, where x is measured from the *front* face of the building, and $c \cong 0.39$. The "constant" c must actually depend on H and L, especially when L is large and the flow reattaches to the sides, but not enough data are available to determine this dependence. The behavior and persistence of the horseshoe vortex require further study, expecially with regard to its interaction with the atmosphere.

Little is known about the elevated vortex pair produced by the interaction of the vertically oriented lee edge vortices with the mean flow near roof level (Fig. 1). Again, Colmer's (1970) field measurements and the smoke releases of Frost and Shahabi (1977) indicate the presence of these vortices, as do wind tunnel experiments (e.g. Counihan, 1971). However, even qualitative descriptions of their behavior and interactions with the mean and cavity flows and with the trailing horseshoe vortex are not yet available.

The three-dimensional extent of the cavity zone is not very well-known. Hosker (1979) examined the along-wind length x_r of the cavity relative to the *rear* face of the building, and suggested the relation

$$\frac{x_r}{H} = \frac{A \cdot (W/H)}{1 + B \cdot (W/H)} \qquad (1)$$

If the flow does *not* reattach to the structure's roof and sides, then A and B are weakly dependent on L/H: $A = -2.0 + 3.7\ (L/H)^{-1/3}$; $B = -0.15 + 0.305\ (L/H)^{-1/3}$. If the flow *does* reattach, A and B are constants: $A = 1.75$, $B = 0.25$. Agreement with data was fair in the case of nonreattachment; the worst case was underpredicted by about 40%. The results were better for the reattached case; predictions were within ±30% of all data, and generally within ±15%. The maximum predicted x_r/H occurs as W/H becomes large; for the reattached case, $(x_r/H)_{\overrightarrow{max}}\ 7$. If reattachment does not take place, the cavity can be longer. The cavity maximum height is equal to H if reattachment to the roof has occurred; if not, values between 1.5 H and 2.5 H have been reported. Roof slope will strongly influence this parameter. The maximum cavity width also depends on the existence of reattachment; if it takes place, then the maximum cavity width = W. Very limited laboratory data suggest that if the flow has not rejoined the building sides, the cavity maximum width is approximately $1.1\ W + 1.7\ W \exp(-0.55\ W/H)$; this relation needs field validation. In order that future work in quantifying the cavity region be facilitated, experimentalists should specify whenever possible the characteristics of the incident flow, such as boundary layer depth, roughness length z_0, and turbulence intensity and scale. Observations of separation and reattachment points would be useful. Field observations of cavity behavior seem to be rather rare and should be encouraged.

The far wake of a building immersed in the atmospheric boundary layer is quite complex, since the presence of potentially persistent longitudinally-oriented vortices generated by the body can strongly influence the flow far downwind. For example, a trailing counter rotating vortex pair can transport higher velocity air from above the wake down into the central portions of the wake, producing a sort of velocity "overshoot"; i.e., the measured speeds in the wake are higher than would be found at that same location if the obstacle were not present. This is particularly apt to be important if the building is at an angle to the incident flow, since the upwind edges of the roof then generate a strong vortex pair quite similar to the wingtip vortices of an airplane (e.g., Hansen and Cermak, 1975). Such phenomena may be significant primarily at distances greater than 10 or 12 H downwind, where momentum deficit-type wake effects have begun to die away. In any event, quantitative evaluation of this far wake region has largely been confined to laboratory studies, and theoretical work is still under development. More field data are vital, especially if the possible significance of vortex-type wakes is to be established. Hansen and Cermak (1975) showed that wake vortex meander due to ambient turbulence results in an apparently weak and rapidly decaying vortex as detected by a fixed sensor, even though the actual vortex may be quite strong and persistent. This means that vortices may be difficult to observe in the field in the wakes of buildings; their presence may have to be inferred from evidence such as velocity overshoot in the wake at large distances downwind (e.g., Colmer, 1970). Furthermore, the vortices can be expected to rise and spread from the wake centerline with increasing downwind distance; hence measurements confined to the surface and centerline may not be very useful.

3. SOURCE UPWIND OF A BUILDING

3.1 Qualitative Description of Concentration Pattern

The above discussion provides some basis for understanding the concentration field near a structure subjected to an isolated incident effluent plume. The rising flow near the upper one-third or so of the upwind face can divert pollutant upward, while the downwash generated by the frontal eddy over the lower two-thirds of the face can be expected to mix the effluent downward, raising concentrations near ground level. Some of the material will be carried off in the trailing horseshoe vortex; pollutant which enters the cavity zone from the sides or from above will be fairly thoroughly mixed in that region, so that concentrations over the rear face of the building will be rather uniform. If the building is sufficiently large in the along-wind direction (i.e., large L), the flow which has separated from the upwind roof edge may reattach, bringing the effluent closer to roof level and possibly inducing high concentrations. If L is small, however, the roof-level flow will not reattach, and the plume will be deflected above roof level; concentrations on the roof then will be similar to those found on the rear face. The overall effect of the building, as viewed from a distance, is to induce a rapid enlargement and dilution of the incident plume; near the building concentrations will be higher than would be observed in its absence, but further downwind the concentrations will be lower.

3.2 Quantitative Results and Calculation Methods

Studies by several investigators have confirmed such expectations. For example, Wilson and Netterville (1978) found a building's influence on the concentration profiles extended only about 1.3 H upwind of its face. Just downwind of the structure they observed quite uniform concentrations and an elevated plume maximum about 1.1 H above the ground (the initial plume height was 1 H). Concentrations were increased locally by factors of two to five over the "no-building" values. As a useful rule of thumb for design of stacks upwind of a building, Wilson and Netterville suggested that the downwash effect generated by a building can be approximated by assuming that the concentration on all building surfaces is uniform and equal to the average of the concentrations expected at roof and at ground level with no building present. They also remarked that, at least for an incident plume released at the height of the "target" building, the concentration at ground level near the building may conservatively be taken equal to the maximum concentration intercepted by any point on the building surface.

Lucas (1972) published a design technique for estimating the required height for an isolated stack so that hazardous or legislated effluent concentration limits are not exceeded on or near the buildings downwind. The method requires a preliminary estimate of required source height using the best available open-terrain model. The height, width, and distance from the source of each downwind structure are also needed. Little explicit use is made of the known flow fields, although their consequences are reflected in the rules of thumb utilized. For example, it is assumed that effluent will not be affected by surrounding structures if the stack is more than 2.5 times higher than all of the neighboring buildings. Also, it is assumed that the turbulent wake of a building can rise to 1.5 times the building width or length (whichever is less), and extend downwind with gradually decreasing effectiveness to 10 building heights or half-widths, whichever is less. The original paper should be consulted for details. Lucas suggested that his technique is appropriate for stacks of 30 m or less, but urged further comparison with experimental data.

A somewhat similar scheme has been suggested by Wilson and Netterville (1976). They remarked that the most important effect of downwind buildings on an incident plume is to mix effluent downward, producing a fairly uniform concentration between roof level and the ground. Hence for design purposes the concentration at roof level should be maintained below the allowable limits, regardless of the location of possibly significant effluent receptors such as building air intakes. To insure this, they introduced the concept of "minimum descent height" Z_m of the maximum permissible concentration χ_m. As indicated in Fig. 3a, Z_m is the height above the surface of the isopleth of maximum allowable concentration at a particular location x. This is a function of the wind speed, since the lower plume boundary will be high above the ground in very light winds, while stronger winds will bend the plume over more sharply. However, strong winds also improve dilution, reducing concentrations in the plume. Consequently the isopleth of allowable concentration will first approach the ground as the wind speed at stack height U_s increases, will reach a height of closest approach (Z_m) to the surface at some critical speed U_A, say, and will then recede from the ground as U_s continues to increase. The procedure is to calculate Z_m by iteration, using an initial guess for the stack height h_s, a standard plume equation with an appropriate plume rise correction, and plot Z_m as a function of downwind distance (Fig. 3b). All stability categories should be considered. On this graph, indicate the locations and heights of downwind structures. If Z_m intersects any existing buildings, or is too low to permit future development of the area, then the stack height should be increased, and the procedure repeated. Alternately, the method may be used by zoning authorities to estimate maximum permissible building heights for new construction downwind of existing stacks. The actual computation is tedious, and best implemented on a computer. Details are available in Wilson and Netterville (1976) and in Hosker (1980).

Wilson and Netterville (1976) made several useful comments about their procedure. First, and perhaps most important, they remarked on the uncertainties inherent in all air pollution calculations. They recommended that the usual "rural" dispersion curves be used, since the extra dilution provided by the larger dispersion parameters characteristic of built-up areas supplies a safety factor. And finally, they noted that the critical atmospheric conditions for air pollution may in practice be either an unstable atmosphere (e.g., looping plumes), or nocturnal plume trapping followed by morning fumigation, whereas the minimum descent height curve is usually closest to the ground under neutrally stable conditions. In other words, the stack designer must consider the full range of possibilities of effluent behavior, and not rely entirely on the technique just outlined. Comparisons of results from the methods of Lucas (1972) and Wilson and Netterville (1976) are needed.

Theoretical work on this upwind source problem by Hunt and his colleagues (Hunt and Mulhearn, 1973; Britter et al., 1976; Puttock and Hunt, 1979; Puttock, 1979), though limited mostly to two-dimensional obstacles with a few exceptions, has generated some useful simplifications. For example, in flows undergoing serious deformation -- as in passing around a building -- the incident plume's cross-sectional area (and concentration) is determined mainly by streamline convergence or divergence, and is not too dependent on local turbulence. If the atmosphere is very stable so that the flow lies in horizontal planes, then the concentration where it strikes the building face will be nearly equal to that predicted for the same location in the absence of the building, as long as the source is at least a few (2 to 10, say) characteristic building dimensions upwind. Behind the obstacle, the theory suggests that, at least for two-dimensional bodies, the concentration in the separated cavity zone is roughly constant and equal to the average along the cavity's separation streamline of the concentration calculated by assuming that streamline to be an impenetrable barrier (Britter et al., 1976; Puttock and Hunt, 1979). The notion of a well-mixed recirculation zone seems to be best fulfilled when the external turbulence level is fairly low. And finally, the "outer" flow, at least for the cases tested thus far, is reasonably well approximated by simple potential flow (Puttock and Hunt, 1979; Puttock, 1979). How well these concepts will carry over to three-dimensional structures remains to be seen. Extensive laboratory and field tests will be needed to validate the results.

4. SOURCE ON OR NEAR A BUILDING

4.1 Qualitative Concentration Patterns

One important flow feature here is the frontal eddy, which induces downwash over the lower two-thirds or so of the building, and then wraps itself around the building ends to trail off downwind. Effluent released from surface vents at mid-height or below on the windward face of a building will therefore be carried to ground level and eventually swirled past the building sides. Some of this material will be advected and diffused into the cavity region, but much of it will be transported past the cavity to the far wake (Hinds, 1969; Meroney and Yang, 1970). On the upper third of the structure, flow will be directed upward and sideward, separating from the windward edges of the roof and sides, and carrying effluent from face vents on the upper portion of the obstacle out and away from other building surfaces. If, however, effluents are released within the separated areas (e.g., on the roof), the resulting ineffective ventilation and flow recirculation can lead to locally high concentrations. If the air intake system for the building is also within the separated zone, a potential health/safety hazard is created, since high concentrations may then be readily circulated throughout the structure. If devices such as heat exchangers are located in these zones of separation and recirculation, they will be prone to self-contamination (plume reingestion), which can seriously degrade their performance and life expectancy.

If L/H for the building is sufficiently large (>1 or 2, say), the edge-separated flow reattaches to the roof and sides, carrying effluent from vents on the windward face closer to the building, but providing better ventilation at outlets located in these reattached areas. Wilson (1977a) noticed that side vents close to the ground produced a somewhat higher effluent concentration at a given distance from the source than did more elevated vents, probably because of the lower wind speed near the ground.

If the building is at an angle to the wind, the flow will remain attached to the most exposed face, but may or may not separate and reattach on the less exposed windward side. Effluent dispersion will vary accordingly. On the roof, a counter-rotating vortex pair will be generated near the upwind corner; these vortices will transport (possibly clean) high-momentum air from above down to roof level. A roof cavity will not occur, and effluents from roof vents will be swept away quickly, keeping rooftop concentrations relatively low.

Behind the structure, the notion of a more or less "closed" recirculating "bubble" bounded by a turbulent shear layer may remain a useful approximation for calculations (Vincent, 1977), although the possibility of direct advection to and from this highly turbulent region must be recognized -- i.e., a "leaky" bubble must be allowed. The cavity seems to be rather nonstationary in atmospheric flows (Frost and Shahabi, 1977), and may even intermittently collapse (Meroney, 1979). Flow adjacent to the lee surface of the obstacle will be more or less upward. Material released within the cavity will be very rapidly mixed throughout its volume, making contamination of any surfaces contacting the flow quite likely (e.g., Meroney and Yang, 1971; Drivas and Shair, 1974; Huber and Snyder, 1976). In particular, for buildings of small L/H ratio, the separated flow from the windward edges of the top and sides becomes part of the cavity; consequently effluents can be expected to appear on the building's roof and sides as well as on its downwind face. Releases somewhat above ground level within the cavity will be mixed more rapidly throughout the region than emissions quite close to the ground (Drivas and Shair, 1974).

The curvature of the turbulent shear layer bounding the cavity from above induces downwash of the "outer" flow, especially near the bubble closure. Material released above the cavity but subject to this flow can be brought close to the ground, producing higher concentrations than would be expected if the building were absent. This has important implications for stack height and placement decisions.

4.2 Effective Stack Height

Successful stack or vent design must take account of building and chimney aerodynamics and effluent buoyancy. Some recent work incorporating such effects is outlined here. The reader is also referred to comprehensive survey reports by Wilson (1976a) and by Meroney (1979).

Some attention must be directed toward the effect of the stack itself on the plume. Sherlock and Stalker (1941) described the production in strong wind of vertically-oriented vortices shed by the stack, as well as a longitudinally-oriented counter rotating vortex pair generated at the stack outlet; these serve to transfer emitted material downward into the stack wake, resulting in a lower effective release height for the material, and producing features such as the commonly observed blackening of chimney tops. Scorer (1968) mentioned the use of flat horizontal disks installed at the stack mouth; if the diameter of these plates is roughly three stack diameters, they prevent effluent from being drawn down into the stack wake cavity, thus preserving the entire useful height of the stack. However, downwash behind a stack is customarily avoided by keeping the efflux momentum at least 50% higher than that of the ambient flow. Sherlock and Stalker (1941) observed downwash distances of only one stack diameter or so for an efflux to ambient velocity ratio of $w_e/U_s = 1$; for $w_e/U_s = 1.5$, practically no downwash was observed. The latter value has been widely used, but should be viewed as a lower bound for design work -- i.e., w_e/U_s should be maintained greater than 1.5 whenever possible to prevent downwash and increase the momentum rise of the effluent plume (e.g., Turner, 1969; ASME, 1973). The effluent density enters the problem as the square root of its ratio to the ambient air density (e.g. Wilson, 1976 a,b), and in most cases should have a rather modest effect on the value of w_e/U_s needed to avoid downwash.

It has been recognized for years that the best way to avoid plume interference by the aerodynamic flow field around an obstruction is to release the material via a tall stack. The familiar "2 and 1/2 times" rule, stating that the stack outlet should be 2.5 times the height of neighboring flow obstacles above the ground, makes use of the rough approximation that the zone of influence of a hill or building generally reaches less than 2.5 H from the surface. Hence material emitted above this height is ejected into a more or less unperturbed flow, permitting reasonable travel distances and significant plume dilution before the effluent diffuses to the ground. In practice, the actual zone of disturbance above the obstacle may be significantly lower than the rule would suggest; this is often true in particular for tall thin structures (e.g., Snyder and Lawson, 1976). To aid stack designers, Briggs (1973) has outlined a tentative procedure for estimating the effective height of emission; the method incorporates corrections for stack downwash, building-induced disturbances to the flow, and buoyancy effects. Model studies of a short wide building and a tall thin building by Snyder and Lawson (1976) have verified that the technique is adequate to avoid obstacle-induced plume downwash. The procedure is as follows:

(1) Compensate for stack-induced downwash by calculating a corrected release height h': if the stack outlet is not vertical, set $h' = h_s$; if it is vertical,

$$h' = h_s + 2\,[(w_e/U_s) - 1.5]d_i \qquad (2)$$

where d_i is the inside diameter of the stack of physical height h_s.

(2) Calculate the release height h'' corrected for building aerodynamic effects. Let ζ_b be the smaller of H or W. If $h' > H + 1.5\zeta_b$, the plume is above the region of building influence; set $h'' = h'$ and go to step (4). If $h' < H$, set

$$h'' = h' - 1.5\,\zeta_b\,; \qquad (3a)$$

if h' is between H and $H + 1.5\,\zeta_b$, set

$$h'' = 2h' - (H + 1.5\zeta_b) \qquad (3b)$$

and go to step (3).

(3) The plume may remain aloft or may be entrained into the wake cavity. If $h'' > 0.5\zeta_b$, the plume remains elevated; go to step (4). On the other hand, if $h'' < 0.5\zeta_b$, the plume is trapped in the cavity, and should be treated as a ground source of initial area ζ_b^2. Go to step (5).

(4) Compensate for plume buoyancy by calculating the effective height of emission h_e. If the effluent is mostly ($\geq$ 98%) air and its temperature is within 2 or 3 C of the ambient, buoyancy is negligible and $h_e = h''$; go to step (5). Otherwise, determine the relative density difference of the plume using

$$\Delta = (1 + 8Q_{H_2O}/Q_T)\,(T_a\,\overline{m}_e)/(T_e\,\overline{m}_a) - 1 \qquad (4)$$

where subscripts e and a denote effluent and ambient, T is absolute temperature, $\bar{m}$ is molecular weight, Q_{H_2O} is the mass of liquid water emitted per unit time, and Q_T is the total mass of effluent per unit time. For most sources the term in Q_{H_2O} is negligible and $\bar{m}_e \simeq \bar{m}_a = 28.9$, so only temperature must be considered. If $\Delta > 0$, the plume is denser than air. A comprehensive treatment of the behavior and resultant concentrations of dense plumes has been given by Hoot *et al.* (1973), which should be consulted. If $\Delta < 0$, the plume is buoyant. Evaluate the buoyancy flux $F = -gV_e\, \Delta/\pi$, where V_e is the effluent volume flux, and calculate the buoyant plume rise Δh using standard procedures (e.g., Briggs, 1970). For large sources, more sophisticated approaches (e.g., Briggs, 1975) may be appropriate. Set $h_e = h'' + \Delta h$, and go to step (5).

(5) Calculate downwind concentrations using standard plume dispersion formulations. Wherever a source height is required, use h_e, the effective height of emission.

4.3 Concentrations on Building Surfaces

If a stack must be placed on a building roof, and if reattachment of the flow to the roof can be expected ($L/H > 1$ or 2), then Wilson's (1979) suggestions may prove useful for avoiding contamination of air intakes or other critical locations on the roof. As mentioned earlier, the roof recirculation zone will reach a maximum height $\cong 0.22R$ a distance $R/2$ from the building's upwind edge; the total cavity length on the roof will be about $0.9R$, where R is the characteristic scale length previously defined. This cavity is bounded above by a turbulent shear layer (zone II in Fig. 2), and above that by the roof wake (zone III). The upper edge of the shear layer, beyond $x=R/2$, is roughly determined by

$$Z_{II}/R \cong 0.27 - 0.1(x/R), \qquad (5)$$

while the roof wake boundary is given by

$$Z_{III}/R \cong 0.28\ (x/R)^{1/3}, \qquad (6)$$

where x is measured from the upwind roof edge. If the stack is high enough that the effluent plume's lower edge remains above the shear layer, contamination of the roof is unlikely; if the plume clears the wake boundary as well, there is no danger of contamination. The required height depends on the rate of plume spread. Wilson suggested that it is generally adequate to draw a line of slope 0.2 in the upwind direction from a point on either the zone II or zone III boundary (depending on the safety margin needed) immediately above the air intake or other critical receptor closest to the lee edge of the roof. The effective stack exit should then lie above this line. This slope corresponds to a standard deviation of wind vertical fluctuation angle σ_ϕ of about 0.1 radians (5.7°), which seems reasonable, although values greater than 0.3 (17.2°) have been observed over open country under some conditions (Lumley and Panofsky, 1964). Whether this value is appropriate to flow over a roof is uncertain.

The sensitivity of local concentrations to variation of parameters such as roof vent height h_s and efflux velocity w_e is of considerable interest. Smith (1975, 1978) conducted detailed field experiments on a small building; he determined concentrations on the roof, sides, and within 5 H downwind for a variety of source and atmospheric conditions. The most prominent feature of his results is the strong dependence of roof concentrations on the ratio h_s/H. For example, a 10% increase in h_s/H caused a 67% decrease in roof concentration, as well as a 25% reduction in concentration on the lee face. Smith also remarked on the varying sensitivity of the results to changes in other parameters; the most significant variables, after h_s/H, are, in order of importance, efflux velocity ratio, angle of wind incidence, turbulence intensity, and the atmospheric stability. He noted that roof concentrations were nearly an order of magnitude lower than the lee face concentration for all stack heights except flush vents, suggesting that roof-mounted air intakes will generally be better than side-mounted inlets whenever the vent height and/or efflux velocity are large enough to keep the emissions from direct contact with the roof.

Wilson (1977b) has prepared a simple argument to demonstrate the sensitivity of roof concentrations to vent height. Effluent is emitted with negligible momentum from a short stack of height $\tilde{h}$ relative to the roof (i.e., $h_s = \tilde{h} + H$). Reattachment of the flow is assumed to occur upwind of the vent, so that the wind seen by the effluent is directed toward the lee edge of the roof. A coordinate system is selected such that $\tilde{x} = 0 = \tilde{y} = \tilde{z}$ at the stack base. The effluent centerline concentration on the roof is then estimated from a standard plume equation; e.g. $\chi \cong Q(\pi\sigma_y\sigma_z U_s)^{-1}\exp(-\tilde{h}^2/2\sigma_z^2)$, so that the nondimensional concentration coefficient is given by $K \equiv \chi U_s A_p/Q \cong A_p(\pi\sigma_y\sigma_z)^{-1}\exp(-\tilde{h}^2/2\sigma_z^2)$. For a flush vent, $\tilde{h} = 0$, so $K_o \cong A_p(\pi\sigma_{yo}\sigma_{zo})^{-1}U_s/U_H$, where the subscript "o" here indicates spreading appropriate to a roof-level release, and A_p is the projected frontal area of the building. After some algebra,

$$K/K_o \approx \exp(-B\, \tilde{h}^2/2\sigma_z{}^2)\ , \qquad (7)$$

where B depends on the receptor location and the wind profile over the building. To examine the relative effectiveness of two stack heights $\tilde{h}_1$ and $\tilde{h}_2$, Eq. (7) gives

$$K_2/K_1 \cong (K_1/K_o)^{(\tilde{h}_2/\tilde{h}_1)^2-1} \qquad (8a)$$

or, in terms of the dilution $D \equiv K_e/K$,

$$D_2/D_1 \approx (D_1/D_o)^{(\tilde{h}_2/\tilde{h}_1)^2-1}\ . \qquad (8b)$$

K_e is the concentration coefficient at the vent exit $= (U_s/w_e)\,(A_p/A_e)$, where A_e is the exit area of the vent. Suppose a short stack of height $\tilde{h}_1$ produces a dilution D_1 of, say, 800 compared to a dilution D_o of, say, 100 produced by a flush vent. If the stack height is increased by 50% to $\tilde{h}_2 = 1.5\,\tilde{h}_1$, then $D_2/D_1 \cong 13.4$, or $D_2 \equiv 10{,}800$. However, if $\tilde{h}_1$ only produces a dilution D_1 of, say, 200 compared to the flush vent value $D_o = 100$, then the same percentage increase in stack height only gives $D_2/D_1 \cong 2.4$, or $D_2 \cong 476$. In other words, an increase in stack height produces really significant decreases in roof-level concentration only if the stack being improved upon was already high enough to cause much more dilution than a flush vent. If values of K_1/K_0 or D_1/D_0 can be obtained from field or laboratory data for a particular building and vent height $\tilde{h}_1$, then the consequences of changes in vent height can be explored using Eq.(8). A handbook of such ratios would be very helpful.

Calculations of dilutions available with flush-mounted roof vents can follow the recommendations of Halitsky (1963a,b) or Wilson (1976b, 1977a). Halitsky (1963a,b) may have been the first to suggest that, for a flush-vent roof source and a receptor anywhere on the roof, side, or rear surfaces of the same building, the dilution is proportional to the square of the shortest distance between the vent and the measurement site. He gave the relation

$$D \cong [\ \tilde{\alpha} + 0.11\ (1 + 0.2\tilde{\alpha})\ (s/\sqrt{A_e})\ ]^2 \qquad (9)$$

where s is the "stretched string" distance between the two points, and $\tilde{\alpha}$ depends on the specific building/vent/ atmospheric configuration. The minimum dilution D_{min} is obtained with $\tilde{\alpha} = 1$. Equation (9) was derived from experiments where $w_e/U_H = 1$; later tests on a more complex building resulted in an expression recommended by ASHRAE (1974) which explicitly utilized the efflux to wind speed ratio:

$$D \cong [\ 4.66 + 0.147\ (s/\sqrt{A_e})]^2\ (U_H/w_e). \qquad (10)$$

Wilson's (1976a,b; 1977a) tests, on the other hand, used low values for w_e/U_H, 0.1 to 0.3 (depending on gas density), which might be more typical of vents with rain deflectors or louvers. He suggested a lower bound to the dilution was

$$D_{min} \cong 0.11\ K_e\ s^2/A_p. \qquad (11)$$

Plots of these equations together with data from several experiments indicate (Wilson, 1976a,b, 1977a; Halitsky, 1979) that Eq. (9), with $\tilde{\alpha} = 1$, is a very conservative lower bound on dilution. On block-like structures it may be an order of magnitude smaller than actual measurements. In this instance Eq. (11) provides somewhat less conservatism, providing a lower bound to the data in most situations. However it overpredicts (Halitsky, 1979) dilutions for block-like buildings at an angle to the wind, and for a nonidealized structure (Clinical Center of the Nat'l. Institute of Health in Bethesda). Equation (10) was specifically developed for this latter site. Overall, it appears that Eq. (11) is probably an adequate lower bound to dilution for most cases, particularly those involving mere nuisance contaminants (e.g., odors). If the effluent is especially troublesome, however, an extra margin of safety may be available using Eq. (9), with $\tilde{\alpha} = 1$. More data, particularly from field tests, would be helpful in validating this guideline.

If the vent is somewhere on the sides of the building, rather than the roof, Wilson (1977a) found that Eq. (11) was still a good lower bound to the dilution, except when the effluent emerged on the front of the structure and subsequently contaminated an air intake close to the ground on the building side. To compensate for this, Wilson suggested that the calculated minimum dilution be further reduced by a factor of five if any potential receptors were located at or below H/5 above the ground.

For more exact work, on-building concentration coefficients K can often be found in the literature (e.g., Halitsky, 1963a,b, 1968; Smith, 1975, 1978; Wilson, 1976a,b; and many others). The results are often but not always for very specific building configurations, such as a particular nuclear reactor. Nevertheless, a great deal of general guidance can be obtained from a careful study of these results and their variation with important parameters such as w_e/U_s and h_s/H. The experimental conditions must be cautiously examined however, particularly for laboratory work, to assess the degree of validity of the data. For example, many early wind tunnel studies were conducted in uniform speed, low turbulence flows not very typical of the atmosphere; consequently important phenomena such as flow reattachment to the building roof or sides did not occur, and the resulting concentration patterns were probably considerably in error. A compendium of these data, together with a discussion of their general applicability and accuracy, would be most useful.

4.4 Concentrations Within the Cavity

As pointed out by Barry (1964) and Meroney (1979), most of the expressions proposed for concentration estimates within the wake cavity take the form

$$\chi \cong \frac{Q}{cUA_p}, \qquad (12)$$

where c is given values between 0.5 and 5.0. The effluent is presumed to be entirely captured in the wake. Evidently K ≅ constant (=1/c), whose value falls between 0.2 and 2.0. A glance at Halitsky's (1963a,b) results, where the concentrations on the lee face should be fairly representative of the cavity as a whole, indicates that these are indeed reasonable values. However, little understanding of the physical mechanisms which generate this result is conveyed; in particular, there is nothing to suggest why the result might vary from one building to another.

Brigg's (1973) recommendation takes some account of the effective stack height and the building size. He defined the concentration coefficient in terms of the smaller building dimension ζ_b : $K \equiv \chi U \zeta_b^2/Q$. Consider the differential stack height h'-H, where h' is given by Eq. (2). If this is greater than $0.35\zeta_b$, then K is generally less than unity in the cavity. However, if h'-H $< 0.35\zeta_b$, K is typically 1.5 and perhaps as large as 3.0.

A slightly more physical hypothesis is based on Vincent's (1977, 1978) ideas. As noted earlier, the cavity "bubble" is not, strictly speaking, a closed zone of recirculation with material transfer possible only across the turbulent shear layer bounding the region; advection into and from the cavity can occur as well. Vincent's (1978) experiments suggest, however, that the bulk of the transfer is conducted by the turbulent exchange processes across the boundary, with only relatively minor contributions from direct advection. Under this somewhat controversial hypothesis, the cavity is modeled to a first approximation as a volume V_c bounded by an active surface of area A_c. Material is emitted into this volume at a rate Q, mixed thoroughly by circulations within the bubble to an average concentration $\bar{\chi}$, and removed in a characteristic time t_d because of turbulent transfer. The mass balance equation can be approximated by

$$\frac{d\bar{\chi}}{dt} + \frac{\bar{\chi}}{t_d} \cong \frac{Q}{V_c} \; . \qquad (13)$$

There are two important possibilities.

Case (1): constant source. At $t = 0$, $\bar{\chi} = 0$ and Q = constant. Then

$$\bar{\chi} \cong \frac{Qt_d}{V_c}\left[1 - e^{-t/t_d}\right] \; . \qquad (14a)$$

Case (2): initial puff. At $t = 0$, $\bar{\chi} = \chi_o$ and $Q = 0$; then

$$\bar{\chi} \cong \bar{\chi}_o e^{-t/t_d} \; . \qquad (14b)$$

The latter result provides a convenient means of evaluating t_d by injecting a known amount of material into the cavity, and plotting the logarithm of the resulting concentration vs. time. This was carried out in field measurements, for example, by Drivas and Shair (1974). The steady-state limit of Eq. (14a) is $\bar{\chi} \cong Qt_d/V_c$, so the average cavity concentration coefficient $\bar{K} \equiv \bar{\chi}\, UA_p/Q$ is

$$\bar{K} \cong A_p U t_d / V_c \; , \qquad (15)$$

which is the ratio of the volume of air intercepted by the building in time t_d to the volume of air in the cavity. A more interesting interpretation is available if V_c is approximated by $A_c x_r$, which may be most appropriate when the flow has reattached to the building roof and sides. Then

$$\bar{K} \sim \frac{t_d}{x_r/U} = \frac{Ut_d/H}{x_r/H} \; . \qquad (16)$$

That is, the cavity concentration coefficient is a measure of the characteristic diffusion time t_d relative to the characteristic "flushing" time of the cavity, x_r/U. Some experiments give values for the nondimensional time $\tau \equiv Ut_d/H$ (e.g., Vincent, 1977, 1978) while x_r/H can be obtained from empirical formulations such as Eq. (1). For a tall thin building or a short wide structure, $\sqrt{A_p}$ or possibly Wilson's (1979) R should probably be substituted for H in the above expressions. Before Eq. (15) or (16) can be used routinely, they must be checked experimentally. For example, direct evaluation of $\bar{K}$ using the limited data of Vincent (1978) and Eq. (16) gives a concentration coefficient roughly a factor of three larger than those cited by Barry (1964); a suitable adjustment factor may be required to compensate for the approximations implicit in the model and for the liberties taken with the turbulent flux parameterization. Direct measurements of $\bar{K}$, t_d, and x_r/H are needed to verify this.

One last comment should be made about locating sources within the cavity: use of a moderately large stack in this region can lead to surprising concentration levels. The work of Huber et al. (1976) on stacks behind a long model ridge illustrates this; in particular, they noted that a stack height ratio $h_s/H = 1.5$ produced the highest observed ground level concentrations near the point of cavity closure. Evidently the effluent was caught in the upper portion of the recirculation pattern and carried directly to the surface. This is one case where a shorter stack may provide lower concentrations than a tall one!

4.5 Concentrations Downwind of the Cavity

For plumes emitted near buildings, but not entrained completely in their wakes, the method of Briggs (1973), outlined above, can be used to estimate an effective stack height for the release. This is then used in a standard Gaussian plume expression for an elevated source.

In cases of complete plume entrainment into the wake cavity, the methods generally used are based on simple modifications to the Gaussian plume equation for a ground source:

$$\chi = \frac{Q}{\pi \sigma_y \sigma_z U} \qquad (17)$$

along the plume centerline.

For example, one can account for the initial rapid mixing of the effluent into the wake by introducing a virtual source some distance upwind of the building. A binormal distribution of material is assumed, with $\sigma_{yo} \cong W/4.3$ and $\sigma_{zo} \cong H/2.15$ (Turner, 1969) representing the dispersion parameters at the building. Next, evaluate the virtual locations x_{yo} and x_{zo} such that $\sigma_y(x_{yo}) = \sigma_{yo}$, $\sigma_z(x_{zo}) = \sigma_{zo}$. Equation (17) is then used, taking $\sigma_y = \sigma_y(x + x_{yo})$, $\sigma_z = \sigma_z(x + x_{zo})$, where x is the distance between the building and the receptor site. Meroney (1979) has recently suggested virtual source locations for a number of likely stack to building height ratios.

A second method, introduced by Gifford (1960), utilizes the idea (attributed to Fuquay) of the wake providing an initial plume dilution proportional to the product of wind speed and projected building area. Ground level concentrations along the plume centerline are then given by

$$\chi = \frac{Q}{(\pi \sigma_y \sigma_z + c A_p)U} \; . \qquad (18)$$

The constant c was estimated on intuitive grounds (Gifford, 1975) to be between 1/2 and 2, with the lower value being the more conservative, as well as agreeing fairly well with assorted test results (see Gifford, 1975, 1976).

A third method introduces "total" diffusion parameters Σ_y and Σ_z(Gifford, 1968, attributed to Davidson):

$$\chi = \frac{Q}{\pi \Sigma_y \Sigma_z U} \; , \qquad (19)$$

where $\Sigma_{y,z} = [\ \sigma_{y,z} + cA_p/\pi\ ]^{1/2}$; the same c values apply as in the previous method.

In many cases the effluent plume will be neither completely elevated nor entrapped by the building wake; instead, a fluctuating partial entrainment occurs, in which portions of the plume are intermittently captured by the wake. Johnson et al. (1975) introduced the so-called "split-h" model to account for the two different effective release heights generated by this intermittency. The model assumes that, in any given hour, a fraction M of the plume will be entrained. The average concentration over an hour is then given by

$$(\chi/Q)_{ave} = M(\chi/Q)_{entr} + (1 - M)\ (\chi/Q)_{elev} \qquad (\ 20\ )$$

The concentration for the entrained portion is calculated from Eq. (19), assuming a Gaussian crosswind distribution, while the elevated concentration is computed from a standard Gaussian formulation with an effective source height estimated from Briggs'(1973) procedure, given above. The entrainment fraction M must depend on parameters such as w_e/U_s, h_s/H, stack location, building geometry, atmospheric stability and turbulence intensity, and wind speed and direction. For the particular case of the Millstone nuclear reactor, Johnson et al. (1975) found the dominant influence was w_e/U_s, and suggested an empirical form for M based on this; their result is probably not appropriate, however, for other sites. Koga and Way (1979) have recently shown the strong dependence of such partial entrainment on h_s/H and vent location, as well as w_e/U_s, for a simple block structure. Experiments such as theirs may eventually produce schemes to estimate M for a particular set of conditions, but this capability is not presently available.

All of these methods have two potential drawbacks, as pointed out by Gifford (1975, 1976). The first, which is not the fault of the models, is that the dispersion parameters generally used do not provide an adequate representation of the plume meander typically observed under light-wind stable conditions. In these circumstances, use of standard expressions or curves for σ_y and σ_z will probably provide significant overestimates of the concentration. Use of site-specific values for the dispersion parameters would alleviate this unwarranted conservatism. The second shortcoming of the methods is inherent in their assumption that the dispersive characteristics of the atmosphere are unaltered by the presence of the building, which serves only to introduce some initial dilution; far downwind the building effect is assumed negligible. This may or may not be true. Wakes under some conditions display a pronounced vortex character. This is particularly apt to occur for buildings at an angle to the wind. These highly persistent vortices typically exhibit rather low swirl velocities, considerable meander, and tend to rise away from the ground with increasing downwind distances. Material entrained within them will not spread in the normal fashion, and may wind up as an elevated plume. Concentrations, especially aloft, may then drop off much more slowly with distance than the simple Gaussian formulation would predict. It is conceivable, then, that the Gaussian predictions will not be conservative under certain circumstances, predicting concentrations lower than those actually present.

Some attention has been devoted recently to development of "improved" expressions for σ_y and σ_z which better account for the effects of an obstacle on the concentration field. Huber and Snyder (1976) examined concentrations in the wake of a model building for various effluent release heights. They observed that emissions between ground level and roof height were dispersed rapidly in both the vertical and crosswind directions, while releases above roof level experienced only enhanced vertical dispersion. On the basis of this work, Huber and Snyder (1976) and Huber (1977, 1979) presented recommended forms based on the concept of spread enhancement generated first by the building cavity and secondly by the decaying turbulence excess produced by the building; for $3 \leq x/H \leq 10$, these are:

$$\sigma_y{}' = 0.7(W/2) + 0.067(x-3H)$$

$$\sigma_z{}' = 0.7H + 0.067(x-3H) \quad , \qquad (\ 21a\ )$$

while for $10 \leq x/H$,

$$\sigma_y{}' = \sigma_y(x + x_{yo})$$

$$\sigma_z{}' = \sigma_z(x + x_{zo}) \qquad (\ 21b\ )$$

Huber uses these in the Gaussian plume model. The virtual source locations x_{yo} and x_{zo} are calculated by requiring that $\sigma_y{}'\ (10W) = 0.7(W/2) + 0.5\ H$, and that $\sigma_z{}'(10H) = 1.2\ H$. For effective stack heights $\leq H$, both horizontal and vertical diffusion enhancement are permitted; for $h_e > H$, only vertical enhancement is used. This method of application of the diffusion parameters is based on the observations of Huber and Snyder (1976).

The agreement of this method with wind tunnel data is fairly good, especially for effective source heights $h_e \leq 1.5\ H$. Agreement with field data seems to range from poor to good on a case by case basis (Huber, 1979). Part of the difficulty is that virtually all of the field work has involved clusters of buildings; it is therefore difficult to determine appropriate single building heights and widths for use in the model. Also, the presence of other buildings between the source and receptor locations in some experiments may have diverted material to other regions. Finally, none of the field studies determined vertical plume height; only ground level measurements of horizontal spread were made. It is impossible to say at this time whether or not the model assumptions (e.g., enhancement in both parameters for low sources, but for the vertical parameter only for higher sources) are verified in the field. Field studies which determine both horizontal and vertical dispersion characteristics are essential.

Concentration measurements behind simple isolated buildings in the field are rather rare, and seem to have been restricted largely to the near-wake region (e.g., Smith, 1975, 1978). The work of Hinds (1969) appears to have been conducted to greater distances downwind than any other, and even it reached only to about $5\ \sqrt{A_p}$. Since the tracer material in his tests was released on the windward side of the building, the near-wake region was not very well mixed (as shown by his isopleth patterns), and the resulting rapid change of concentration with distance is probably not at all typical of that behind a building with a cavity-entrained plume.

Meroney and Yang (1970, 1971) examined the concentrations behind a cubical model to as much as 50 H downwind. For ground level releases, their concentration coefficient K decreased roughly as x to the -0.6 to -0.7 power. This is a much slower decay rate than would be estimated from the open terrain Gaussian model, which would predict powers of -1.3 to -1.6, depending on stability.

Symes and Meroney (1970) reported concentration measurements behind a short cylinder and a sequence of cone frustrums of various tapers. The isopleth patterns varied somewhat with the body shape, but the centerline ground-level concentrations beyond $x/H \cong 10$ were all quite similar, decaying as x to the -0.8 to -1.0 power. The measurements extend to about 35 H downwind. Huber and Snyder's (1976) data indicates that between 10 H and 30 H downwind of an effectively ground-level source, the concentration falls off as $x^{-0.8}$; beyond 30 H, however, their wake-enhanced Gaussian model approaches a -1.3 power decay. Some of these apparent discrepancies may be due to the presence of large scale turbulence and consequent plume meander in the field, an effect usually impossible to duplicate in a wind tunnel. This is discussed at some length below. In any case, field data at large x/H for isolated buildings would be most welcome.

5. SOURCE DOWNWIND OF A BUILDING

5.1 Qualitative Description of Concentration

Relatively little guidance is available for estimating the concentrations produced by a source located some distance downwind of an isolated building, so this section is necessarily speculative. Suppose the source is located well back in the wake cavity. If it is a low or ground level vent, then the effluent will travel toward the building's lee face, and be well stirred throughout the cavity in the process. A higher stack at this location may have its effluent transported fairly directly to the surface by the downward curving shear layer bounding the cavity. If the cavity is fluctuating or collapsing, these building wake effects will be intermittently important. If the source is off the wake centerline, and if the horseshoe vortex system does not break down rapidly in the ambient turbulence, then effluent may be entrained in one of the vortices and be carried off downwind, possibly rising from the surface. If an elevated vortex pair is important, as may be true for buildings at an angle to the wind, then vortex entrainment and transport may be observed for an elevated source as well. In this case one might find effluent reaching the surface far downwind of the structure. Similar arguments may be appropriate for sources downwind of the cavity closure. One expects the enhancement of turbulent diffusion induced by the building wake to be the dominant effect, so that effluent from an elevated source is mixed to the ground more rapidly, while concentrations far downwind are lower than if the building were not present. If organized vorticity is present in the wake, however, it may be able to transfer rather concentrated material aloft to the surface at large distances downwind.

5.2 Calculations

Briggs (1973) suggested that his method of incorporating building wake effects into the effective stack height should be confined to cases where the stack is within $3\zeta_b$ downwind of the building -- i.e., within or near the end of the cavity zone. His recommended model for computing downwind concentrations if the effective stack height is zero is essentially Eq. (18) with c = 1. If the plume remains aloft, he used $\chi = 0$ for $\sigma_z < \sqrt{2/\pi}\, h_e$, while for $\sigma_z > \sqrt{2/\pi}\, h_e$,

$$\chi \cong \frac{Q}{\pi\, \sigma_y\, (\sigma_z + \sqrt{2/\pi}\, h_e) U} \, . \qquad (22)$$

That is, dispersion is enhanced only in the vertical for an elevated release in the presence of the building. No contribution is allowed if the source is more than $3\zeta_b$ downwind.

Huber (1977, 1979) commented that the Huber and Snyder dispersion enhancement model is best applied to sources within 2 H of a building, with $h_e < H + 1.5\zeta_b$. No recommendation was made for more distant sources.

In fact, there seems to have been little study of the concentration field produced by a source located downwind of the cavity zone of a building or building cluster, despite its practical importance. Barrett *et al*. (1978) studied the case of a single model building at an angle to the wind for a range of stack heights and downwind locations. They remarked that examination of the velocity and turbulence profiles behind the structure would suggest that the wake is virtually dissipated by $x/H \cong 8$, yet the effects of the building on the concentration field can be observed out to at least 24 H downwind. They attributed this to downwash along the wake centerline induced by a persistent trailing horseshoe vortex. Figure 4 shows their maximum groundlevel concentration coefficient $K_{max} \equiv \chi_{max} U_H H^2/Q$ as a function of nondimensional stack height h_s/H, for a range of stack locations downwind. Further study and field validation are needed for a variety of building shapes and orientations.

6. CLUSTERS

6.1 Local Effects

Pollutant dispersion of an incident plume by a cluster of buildings can only be described in rather general terms because of the enormous variations in flow patterns possible, even for relatively simple configurations. In particular, gross plume deflections near long walls, "jetting" effects near gaps, and a strong sensitivity to wind direction can be expected. Wind tunnel and/or field tests are virtually mandatory for any realistic assessment of a particular site. In practice, nearly all field studies of diffusion in building wakes actually have involved such clusters, typically for major industrial plants such as power stations or reactor sites. It is consequently often difficult to separate general trends from special effects generated by site-dependent flow interactions between and behind the obstacles. Such complicated flows may be at least partially responsible for the observed variations in far wake behavior in particular.

Cagnetti (1975) described a field study of a source upwind of a comparatively simple reactor complex which illustrates some of the problems. For several tests, his patterns of constant K did not resemble in any way the characteristic closed "footprint' of plumes in open terrain. Instead, the effluent seemed to be largely deflected to the sides by the upwind building surfaces, and the K values were fairly small. After a modest shift (~30°) in wind direction, however, the bulk of the flow was directed toward a gap between buildings; the isopleth pattern then was quite reminiscent of a classical plume, while the K values were about an order of magnitude larger behind the buildings.

Much of the field work accomplished in the near wake region of building clusters is probably applicable to single building wakes. For example, Abbey (1976) and Start et al. (1977) described the rather similar strong vertical mixing of smoke within the cavity of two dissimilar reactor buildings. During low windspeed stable conditions in particular, the wake acted to convert a ground level source into an effectively elevated one, so that plume concentration maxima appeared aloft downwind. A number of investigators have noted "critical" windspeeds above which stack or roof emitted effluents were downwashed into the wake cavity (e.g., Davies and Moore, 1964; Martin, 1965; Munn and Cole, 1967a,b; Rodliffe and Fraser, 1971). A bimodal state is often postulated for such cases, such that the plume remains aloft for low wind speeds, but is fully entrained for others (e.g., Johnson et al., 1975; Lawson, 1967).

Downwind of the cavity behind the structure serving as an effluent release point, but within or close to the building cluster as a whole, observed concentration isopleths display widely varied behavior, expecially when other large obstacles are present in the cluster. For example the data of Start et al. (1977) showed a great sensitivity to wind direction and source location, and highly individual effects such as cavity concentration maxima and zones of local "jetting" between structures were apparent. Strong site dependence is inevitable for these situations. Mathematical modeling of near-field behavior for such sites on an episode by episode basis will be extremely difficult, if not impossible. On the other hand, Dickson et al. (1969) presented average concentration isopleths for lapse and inversion conditions, for an ensemble of fifteen tests. These K patterns were much like those of a classical isolated plume, although there was significant initial broadening. Simple modeling efforts may well be effective for predicting such ensemble average behavior (e.g., Huber's (1979) work).

Halitsky (1977) developed a moderately complicated mathematical model for a building complex and applied it to the site studied by Dickson et al. (1969). The model predicts that the wake boundary grows in the horizontal and vertical as $x^{1/3}$, and that the mean velocity defect and turbulence intensity excess both decay as $x^{-2/3}$. These predictions do not agree with those made by Counihan et al. (1974) for the wake behind an isolated building. However Halitsky's (1977) mean velocity defect prediction appeared to agree well with Dickson et al's (1969) data. By adjustment of a free parameter related to the size of the complex, excellent agreement of calculated K isopleths with the data was also obtained. One interesting point is that the crosswind concentration distribution, matched to Dickson et al's (1969) data, was taken to be parabolic rather than Gaussian, at least close to the reactor complex. Further investigation of this would be welcome.

6.2 Far Wake Effects

Abbey (1976) summarized the observed decrease of concentration with downwind distance in the far wakes of building clusters. Figures (5a,b) show the results of field tests and wind tunnel studies, respectively, for neutral and unstable atmospheric conditions, while Fig. (5c) combines the field and laboratory data for stable conditions. Beyond a few hundred meters downwind, a power-law dependence on distance is evident. For field tests under neutral and unstable conditions, the power law exponent ranges between -1.1 and -1.8. The wind tunnel tests, on the other hand, have exponents between -0.3 and -1.9. A laboratory study by Hatcher and Meroney (1977), not shown in these figures, found a decay rate of $x^{-1.8}$ with little dependence on stability. Halitsky's (1977) mathematical model for wake concentration seemed to decrease as $x^{-1.2}$, agreeing quite well with Dickson et al's (1969) field value of $K \equiv \chi U A_p / Q \cong 5(x/H)^{-1.3}$ at large values of x, where H was the containment vessel height. Under stable conditions, the field data in Fig. (5c) behave as $x^{-1.3}$ or so, while both laboratory results decay much more slowly, as $x^{-0.7}$. The "Islitzer" result in this figure decays very rapidly, as x^{-3}, but Abbey (1976) remarked that it was due to only a single test, and therefore may not be too reliable. Cagnetti's (1975) far wake K data decreased as $x^{-1.7}$.

The reason for these discrepancies between wind tunnel and field far wake data is unknown, although some speculation is possible. For example, under stable light wind conditions, wind "meander" is well known as an important contributor to plume dilution. Start et al. (1977) noted wind direction fluctuations so large that lateral plume spreading induced by the building complex under study was nearly obscured. Such meander is virtually impossible to duplicate in the laboratory, a fact which may help to account for the relatively slow decrease of concentration with distance that is observed in the wind tunnel studies. It may also be true that persistent vortex-containing wakes are more significant in the laboratory than in the field; if so, the resulting slow spread of the wake would inhibit plume dispersion, producing a lower concentration decay rate that has been observed in field experiments. The possible enhanced importance of such vortex phenomena in laboratory work could be due to many factors, such as lower Reynolds number, incomplete simulation of the turbulence properties of the incident wind, and improper scaling of the model buildings relative to the incident boundary depth, turbulence scales, and other variables. Additional work to resolve this dilemma is needed; in the meantime, wind tunnel data on far wake concentration patterns behind building clusters must be viewed with caution.

7. SUMMARY

The last decade has seen the development of a degree of largely empirical predictive capability for estimating concentrations near buildings. Many of the schemes are based, however, on a limited amount of data, usually collected in the laboratory. Other methods are essentially logical procedures rooted in a general understanding of the major pertinent phenomena. There are some conflicts in the techniques, and practically none have been validated in the field. Existing field data are in most cases difficult to interpret and compare to simple predictive models because the presence of multiple buildings in most previous tests confuses the issue. In view of recent advances in our knowledge of near-building flows, it may now be appropriate to conduct additional field tests of isolated simple structures. Measurements should be made to large distances ($x/H > 20$ or 30), and should not be confined to the ground-level wake centerline. Remote-sensing devices may be useful. Flow and concentration data from carefully planned and executed field tests conducted under predominantly neutral atmospheric conditions would be extremely useful for validating present theoretical and laboratory results, and improving our understanding of this very complex problem. In the meantime, though some general guidance is available, individual tests in the wind tunnel and/or field remain the best approach to determining effluent concentrations, especially in critical cases.

8. ACKNOWLEDGEMENT

Most of this work was accomplished with the support of the U. S. Nuclear Regulatory Commission; some portions were completed under agreements between NOAA and the U. S. Department of Energy. The author is grateful to Drs. G. A. Briggs and K. S. Rao for their helpful comments.

9. REFERENCES

Abbey, R. F., Jr., 1976: Concentration measurements downwind of buildings: previous and current experiments. In Preprints of Third Symposium on Atmospheric Turbulence, Diffusion and Air Quality, Oct. 19-22, Raleigh, NC (Amer. Met. Soc., Boston, MA), pp. 247-254.

ASHRAE, 1974: Handbook and Product Directory, 1974, Applications, Ch. 15 "Laboratories" (Amer. Soc. Heating, Refrig., and Air-Cond. Engineers, New York), pp. 15.9 ff.

ASME, 1973: Recommended Guide for the Prediction of Airborne Effluents, second edition (Amer. Soc. Mech. Engineers, New York), xii and 85 pp.

Baines, W. D., 1963: Effects of velocity distribution on wind loads and flow patterns on buildings. In Proc. of Symposium on Wind Effects on Buildings and Structures, National Physical Laboratory, Teddington, Middlesex, England, June 26-28, (Her Majesty's Stationery Office, London, 1965), Vol. I, pp. 197-225.

Barrett, C. F., D. J. Hall, and A. C. Simmonds, 1978: Dispersion from Chimneys Downwind of Cubical Buildings -- a wind-tunnel study. In Proc. of Ninth International Technical Meeting on Air Pollution Modeling and its Application, Toronto, Canada, Aug. 28-31 (North Atlantic Treaty Organization, CCMS), pp. 37-44.

Barry, P. J., 1964: Estimation of downwind concentration of airborne effluents discharged in the neighbourhood of buildings. AECL Report no. 2043 (Atomic Energy of Canada, Ltd., Chalk River, Ontario), 16 pp.

Briggs, G. A., 1970: Some recent analyses of plume rise observations. In Proc. of Second International Clean Air Congress, Washington, D.C., Dec. 6-11, H.M. Englund and W. T. Beery, eds. (Academic Press, New York, 1971), pp. 1029-1032.

Briggs, G. A., 1973: Diffusion estimation for small emissions. In Atmospheric Turbulence and Diffusion Laboratory 1973 Annual Report (NOAA/ATDL, Oak Ridge, TN, publication no. ATDL-106), pp. 83-145.

Briggs, G. A., 1975: Plume rise predictions. Chapter 3 in Lectures on Air Pollution and Environmental Impact Analyses, Boston, Sept. 29-Oct. 3 (Amer. Met. Soc., Boston, MA), pp. 59-111.

Britter, R. E., J. C. R. Hunt, and J. S. Puttock, 1976: Predicting pollution concentrations near buildings and hills. Presented at Institution of Measurements and Control Conference on Systems and Models in Air and Water Pollution, London, Sept. 22-24, pp. 7-1 to 7-15.

Cagnetti, P., 1975: Downwind concentrations of an airborne tracer released in the neighbourhood of a building. Atmos. Environment, 9, no. 8, pp. 739-747.

Colmer, M. J., 1970: Some Full-Scale Measurements of the Flow in the Wake of a Hanger. Royal Aircraft Establishment/Aeronautical Research Council C.P. no. 1166, replacing RAE Tech. Report 70202-ARC32863 (Her Majesty's Stationery Office, London, 1971), 28 pp.

Corke, T. C., and H. M. Nagib, 1976: Sensitivity of Flow Around and Pressures on a Building Model to Changes in Simulated Atmospheric Surface Layer Characteristics. Illinois Institute of Technology Fluids and Heat Transfer Report R76-1 (May), (IIT, Mechanics & Mech. & Aerospace Engineering Dept., Chicago, IL), xxiii and 268 pp.

Counihan, J., 1971: An Experimental Investigation of the Wake Behind a Two Dimensional Block and Behind a Cube in a Simulated Boundary Layer Flow. Central Electricity Research Laboratories note no. RD/L/N 115/71 (CERL, Leatherhead, Surrey, England), 45 pp.

Counihan, J., J. C. R. Hunt, and P. S. Jackson, 1974: Wakes behind two-dimensional surface obstacles in turbulent boundary layers. J. Fluid Mech., 64, part 3, pp. 529-563.

Davies, P. O. A. L., and D. J. Moore, 1964: Experiments on the behavior of effluent emitted from stacks at or near the roof level of tall reactor buildings. *Internat. J. Air Water Pollution*, 8, no. 10, pp. 515-533.

Dickson, C. R., G. E. Start, and E. H. Markee, Jr., 1969: Aerodynamic effects of the EBR-II reactor complex on effluent concentration. *Nuclear Safety*, 10, no. 3, pp. 228-242.

Drivas, P. J., and F. H. Shair, 1974: Probing the air flow within the wake downwind of a building by means of a tracer technique. *Atmos. Environment*, 8, pp. 1165- 1175.

Frost, W., and A. M. Shahabi, 1977: A Field Study of Wind Over a Simulated Block Building, NASA report no. CR-2804 (avail. NTIS, Springfield, VA), xii and 119 pp.

Gifford, F. A., 1960: Atmospheric dispersion calculations using the generalized Gaussian plume model. *Nuclear Safety*, 2, no. 2, pp. 56-59.

Gifford, F. A., 1968: An outline of theories of diffusion in the lower layers of the atmosphere. Chapter 3 in *Meteorology and Atomic Energy-1968*, D. Slade, editor (U.S. AEC no. TID-24190, avail. NTIS, Springfield, VA), pp. 65-116.

Gifford, F. A., 1975: Atmospheric dispersion models for environmental pollution applications. Chapter 2 in *Lectures on Air Pollution and Environmental Impact Analyses*, Boston, Sept. 29-Oct. 3 (Amer. Meteorol. Soc., Boston, MA), pp. 35-58.

Gifford, F. A., 1976: Turbulent diffusion-typing schemes; a review. *Nuclear Safety*, 17, no. 1, pp. 68-86.

Halitsky, J., 1963a: Gas Diffusion Near Buildings. N.Y.U. Meteorol. and Oceanog. Geophys. Sciences Lab. Rep. no. 63-3 (New York University, College of Engineering, New York, NY).

Halitsky, J., 1963b: Gas diffusion near buildings, *ASHRAE Trans.* 69, paper no. 1855, pp. 464-485.

Halitsky, J., 1968: Gas diffusion near buildings. Ch. 5-5 in *Meteorology and Atomic Energy-1968*, ed. by D. H. Slade (U.S. AEC, TID-24190, avail. NTIS, Springfield, VA), pp. 221-255.

Halitsky, J., 1977: Wake and dispersion models for the EBR-II building complex. *Atmos. Environment*, 11, no. 7, pp. 577-596.

Halitsky, J., 1979: Private communication.

Halitsky, J., J. Golden, P. Halpern, and P. Wu, 1963: Wind Tunnel Tests of Gas Diffusion From a Leak in the Shell of a Nuclear Power Reactor and From a Nearby Stack. N.Y.U. Meteorol. and Oceanog. Geophys. Sciences Lab. Rep. no. 63-2 (New York University, College of Engineering, New York, NY).

Hanna, S. R., G. A. Briggs, and R. P. Hosker, Jr., 1980: A Handbook on Atmospheric Diffusion (U.S. Dept. of Energy Tech. Information Center, Oak Ridge, TN -- in preparation).

Hansen, A. C., and J. E. Cermak, 1975: Vortex-Containing Wakes of Surface Obstacles. Colorado State Univ. Fluid Dynamics and Diffusion Lab. report CER75-76ACH-JEC16 (Col. State Univ., Ft. Collins), xiv and 163 pp.

Hatcher, R. V., and R. N. Meroney, 1977: Dispersion in the wake of a model industrial complex. In Preprints of *Joint Conference on Applications of Air Pollution Meteorclogy*, Salt Lake City, Nov. 29-Dec. 2 (Amer. Meteorol. Soc., Boston, MA), pp. 343-346.

Hinds, W. T., 1969: Peak-to-mean concentration ratios from ground-level sources in building wakes. *Atmos. Environment*, 3, no. 2, pp. 145-156.

Hoot, T. B., R. N. Meroney, and J. A. Peterka, 1973: Wind Tunnel Tests of Negatively Buoyant Plumes. Colorado State Univ. Fluid Dynamics and Diffusion Lab. report CER73-74TGH-RNM-JAP-13 (Col. State Univ., Ft. Collins).

Hosker, R. P., 1979: Empirical estimation of wake cavity size behind block-type structures. In Preprints of Fourth *Symposium on Turbulence Diffusion, and Air Pollution*, Reno, NV, Jan. 15-18 (Amer. Meteorol. Soc., Boston, MA, 1979), pp. 603-609.

Hosker, R. P., Jr., 1980: Flow and dispersion near obstacles. Chapter 7 in *Atmospheric Sciences and Power Production*, D. Randerson, editor (U.S. Dept. of Energy Tech. Information Center, Oak Ridge, TN -- in preparation).

Huber, A. H., 1977: Incorporating building/terrain wake effects on stack effluents. In *Preprints of Joint Conference on Applications of Air Pollution Meteorology*, Salt Lake City, Nov. 29-Dec. 2 (Amer. Meteorol. Soc., Boston, MA), pp. 353-356.

Huber, A. H., 1979: An evaluation of obstacle wake effects on plume dispersion. Presented at *Fourth Symposium on Turbulence, Diffusion, and Air Pollution*, Reno, NV, Jan. 15-18, (Amer. Meteorol. Soc., Boston, MA), 8 pp.

Huber, A. H., and W. H. Snyder, 1976: Building wake effects on short stack effluents. In *Preprints of Third Symposium on Atmospheric Turbulence, Diffusion and Air Quality*, Raleigh, NC, Oct. 19-22 (Amer. Meteorol. Soc., Boston, MA), pp. 235-242.

Huber, A. H., W. H. Snyder, R. S. Thompson, and R. E. Lawson, Jr., 1976: Stack Placement in the Lee of a Mountain Ridge. U.S. E.P.A. Environmental Sciences Research Laboratory report no. EPA-600/4-76-047 (September), (avail. NTIS, Springfield, VA), x and 44 pp.

Hunt, J. C. R., C. J. Abell, J. A. Peterka, and H. Woo, 1978: Kinematical studies of the flows around free or surface mounted obstacles; applying topology to flow visualization. *J. Fluid Mech.*, 86, part 1, pp. 179-200.

Hunt, J. C. R., and P. J. Mulhearn, 1973: Turbulent dispersion from sources near two-dimensional obstacles. *J. Fluid Mech.*, *61*, part 2, pp. 245-274.

Hunt, J. C. R., 1975: Turbulent velocities near and fluctuating surface pressures on structures in turbulent winds. In *Proc. of the Fourth Internat. Conf. on Wind Effects on Buildings and Structures*, Heathrow (London), U.S., Sept. 8-12 (Cambridge Univ. Press, London, 1977), pp. 309-320.

Islitzer, N. F., 1965: Aerodynamic Effects of Large Reactor Complexes Upon Atmospheric Turbulence and Diffusion. U.S. AEC Nat. Reactor Test. Station report IDO-124041 (U.S. Weather Bureau, Idaho Falls, ID), 15 pp.

Jensen, M., and N. Frank, 1963: Model-Scale Tests in Turbulent Wind, Part I (The Danish Tech. Press, Copenhagen) viii and 97 pp.

Johnson, W. B., E. Shelar, R. E. Ruff, H. B. Singh, and L. Salas, 1975: Gas Tracer Study of Roof-Vent Effluent Diffusion at Millstone Nuclear Power Station. Atomic Industrial Forum report AIF/NESP-007b (Stanford Research Institute, Palo Alto, CA).

Koga, D. J., and J. L. Way, 1979: Effects of Stack Height and Position on the Dispersion of Pollutants in Building Wakes. I.I.T. Fluids and Heat Transfer Rep. no. R79-2 (Illinois Inst. of Technology, MMAE Dept., Chicago, IL) xxiii and 238 pp.

Lawson, T. V., 1967: Discussion of "Turbulence and diffusion in the wake of a building." *Atmos. Environment* *1*, no. 2, pp. 177-181.

Lemberg, R., 1973: *On the Wakes Behind Bluff Bodies in a Turbulent Boundary Layer*. Ph.D. thesis and Boundary Layer Wind Tunnel report no. BLWT-3-73 (University of Western Ontario, London, Ontario, Canada), xiv and 159 pp.

Lucas, D. H., 1972: Choosing chimney heights in the presence of buildings. In *Proc. Internat. Clean Air Conf.*, Melbourne Univ., Australia, May (Clean Air Society of Australia and New Zealand, 1972), pp. 47-52.

Lumley, J. L., and H. A. Panofsky, 1964: *The Structure of Atmospheric Turbulence* (John Wiley-Interscience, New York, NY), xii and 239 pp.

Martin, J. E., 1965: *The Correlation of Wind Tunnel and Field Measurements of Gas Diffusion Using Krypton-85 As a Tracer*. Ph.D. Thesis, and report no. MMPP-272 of Michigan Memorial Phoenix Project (Univ. of Michigan, Ann Arbor), x and 130 pp.

Meroney, R. N., 1979: Turbulent diffusion near buildings. Chapter 11 in *Engineering Meteorology*, E. J. Plate, editor (to be published).

Meroney, R. N., and B. T. Yang, 1970: Gaseous plume dispersion about isolated structures of simple geometry. Presented at Second International Clean Air Congress of the Internat. Union of the Air Pollution Prevention Association, Washington, DC, Dec. 6-11, 32 pp.

Meroney, R. N., and B. T. Yang, 1971: Wind-Tunnel Study on Gaseous Mixing Due to Various Stack Heights and Injection Rates Above an Isolated Structure. Colorado State Univ. Fluid Dynamics and Diffusion Lab. report CER71-72RNM-BTY16 (Col. State Univ., Ft. Collins), v and 42 pp.

Morkovin, M. V., 1972: An approach to flow engineering via funcational flow modules. In *Beiträge z. Strömungsmechanik*, A. Walz 65th Anniversary Volume, Deutsche Luft-und Raumfahrt Forschungsvericht no. 72-27.

Munn, R. E., and A. F. W. Cole, 1967a: Turbulence and diffusion in the wake of a building. *Atmos. Environment*, *1*, no. 1, pp. 33-43.

Munn, R. E., and A. F. W. Cole, 1967b: Some strong-wind downwash diffusion measurements at Douglas Point, Ontario, Canada. *Atmos. Environment*, *1*, no. 5, pp. 601-604.

Norman, R. S., 1972: *On Obstacle Generated Secondary Flows in Laminar Boundary Layers and Transition to Turbulence.* Ph.D. thesis, Illinois Institute of Technology (Mech. Engineering Dept.), xvi and 151 pp.

Ogawa, Y., 1973: *Effects of Building and Thermal Boundary Layer On Diffusion.* Ph.D. thesis, Hokkaido Univ., Sapporo, Japan (Dept. of City Environmental Engineering), xv and 327 pp.

Penwarden, A. D., and A. F. E., Wise, 1975: Wind Environment Around Buildings. Building Research Establishment Report (Her Majesty's Stationery Office, London), vii and 52 pp.

Puttock, J. S., 1979: Turbulent diffusion from sources near obstacles with separated wakes - Part II. Concentration measurements near a circular cylinder in uniform flow. *Atmos. Environment*, *13*, no. 1, pp.15-22.

Puttock, J. S. and J. C. R. Hunt, 1979: Turbulent diffusion from sources near obstacles with separated wakes - Part I. An eddy diffusivity model. *Atmos. Environment*, *13*, no.1, pp. 1-13.

Rodliffe, R. S., and A. J. Fraser, 1971: Measurements on the release of gaseous activity from a short stack. *Atmos. Environment*, *5*, no. 4, pp. 193-208.

Scorer, R. S., 1968: *Air Pollution* (Pergamon Press, Oxford, London, New York, Toronto, and Paris), xiv and 151 pp.

Sherlock, R. H., and E. A. Stalker, 1941: A Study of Flow Phenomena in the Wake of Smokestacks. Univ. of Michigan Dept. of Engineering Research Bulletin No. 29 (Univ. of Michigan Press, Ann Arbor), vi and 49 pp.

Smith, D. G., 1975: Influence of meteorological factors upon effluent concentrations on and near buildings with short stacks. Presented at 68th Annual Meeting of Air Poll. Control Assoc., Boston, June 15-20, paper #75-26.2, 25 pp.

Smith, D. G., 1978: Turbulent Dispersion Around a Building in the Natural Wind and Effluent Reentry. Ph.D. thesis, Harvard Univ. (School of Public Health, Boston, MA), xvi and 150 pp.

Snyder, W.H., and R. E. Lawson, Jr., 1976: Determination of Height for Stack Near Building, U.S. E.P.A. Environmental Monitoring Series report no. EPA-600/4- 76-001 (avail NTIS, Springfield, VA), iii and 30 pp.

Start, G. E., J. H. Cate, C. R. Dickson, N. R. Ricks, G. R. Ackermann, and J. F. Sagendorf, 1977: Rancho Seco Building Wake Effects on Atmospheric Diffusion. NOAA Tech. Memorandum ERL-ARL-69 (NOAA/ARL, Idaho Falls, ID), iv and 185 pp.

Symes, C. R., and R. N. Meroney, 1970: Cone Frustrums in a Shear Layer. U.S. AEC report COO-2053-4, Colorado State Univ. Fluid Dynamics and Diffusion Lab. report CER70-71CRS-RNM11 (Colo. State Univ., Ft. Collins), xi and 131 pp.

Turner, D. B., 1969: Workbook of Atmospheric Dispersion Estimates. (U.S. Dept. of HEW Public Health Service, Publication no. 999-AP-26), viii and 84 pp.

U.S. E.P.A., 1978: Guideline on Air Quality Models. Environmental Protection Agency Office of Air Qual. Planning and Standards report no. EPA-450/2-78-027 (OAQPS no. 1.2-080), April (U.S. EPA, Research Triangle Park, NC), viii and 84 pp.

Vincent, J. H., 1977: Model experiments on the nature of air pollution transport near buildings. Atmos. Environment, 11, no. 8, pp. 765-774.

Vincent, J. H., 1978: Scalar transport in the near aerodynamic wakes of surface mounted cubes. Atmos. Environment, 12, no. 6/7, pp. 1319-1322.

Wilson, D. J., 1976a: Contamination of Building Vents. Univ. of Alberta Dept. of Mech. Engineering report no. 1 (Univ. of Alberta, Edmonton, Alberta, Canada), iv and 126 pp.

Wilson, D. J., 1976b: Contamination of air intakes from roof exhaust vents. ASHRAE Trans. 82, part 1, 1024-1038.

Wilson, D. J., 1977a: Dilution of exhaust gases from building surface vents. ASHRAE Trans. 83, part 1, 168-176.

Wilson, D. J., 1977b: Effect of vent stack height and exit velocity on exhaust gas dilution. ASHRAE Trans 83, part 1, 157-167.

Wilson, D. J., 1979: Flow patterns over flat-roofed buildings and application to exhaust stack design. ASHRAE Trans. 85, part 2, 284-295.

Wilson, D. J., and D. D. J. Netterville, 1976: Influence of downwind high-rise buildings on stack design. J. Air Poll. Control Assoc., 26, no. 10, pp. 976-980.

Wilson, D. J., and D. D. J. Netterville, 1978: Interaction of a roof-level plume with a downwind building. Atmos. Environment, 12, no. 5, pp. 1051-1059.

Woo, H. G. C., J. A. Peterka, and J. E. Cermak, 1977: Wind-Tunnel Measurements in the Wakes of Structures. NASA contractor report NASA CR-2806 (March) (avail. NTIS, Springfield, VA), xvi and 226 pp.

Yang, B. T., and R. N. Meroney, 1970: Gaseous Dispersion Into Stratified Building Wakes. Colorado State Univ. Fluid Dynamics and Diffusion Lab. report CER70-71BTY- RNM-8 (Col. State Univ., Ft. Collins), xi and 103 pp.

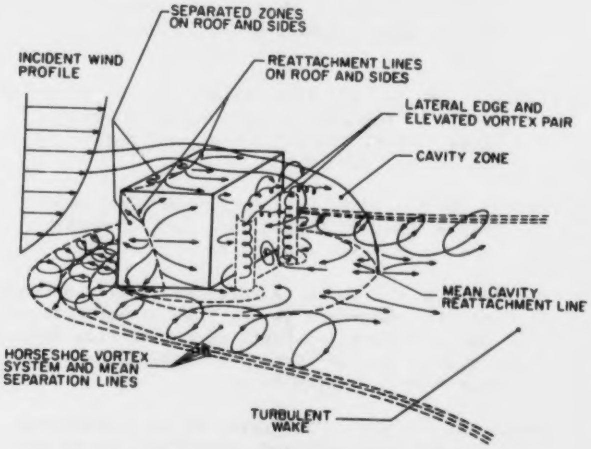

SEPARATED ZONES ON ROOF AND SIDES
INCIDENT WIND PROFILE
REATTACHMENT LINES ON ROOF AND SIDES
LATERAL EDGE AND ELEVATED VORTEX PAIR
CAVITY ZONE
MEAN CAVITY REATTACHMENT LINE
HORSESHOE VORTEX SYSTEM AND MEAN SEPARATION LINES
TURBULENT WAKE

FIGURE 1. RECENT MODEL OF FLOW NEAR BUILDING (based on Woo et al., 1977; Hunt et al., 1978).

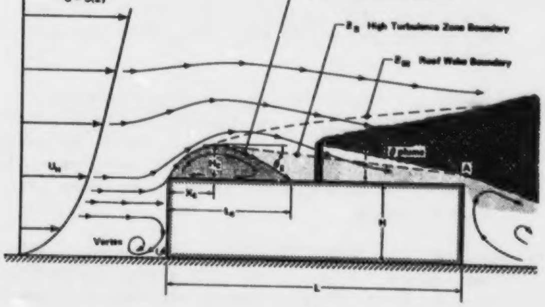
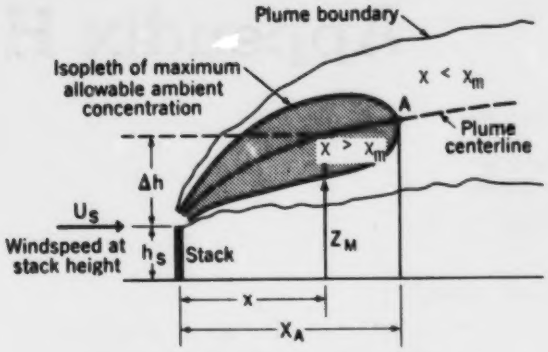
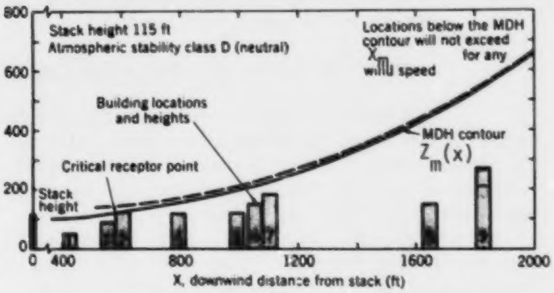

Upwind Velocity Profile
U = U(z)
Z_I Recirculation Cavity Boundary
Z_{II} High Turbulence Zone Boundary
Z_{III} Roof Wake Boundary
plume
A
U_H
X_c
L_c
H
Vortex
L

FIGURE 2. FLOW OVER LONG FLAT BUILDING ROOF FOR WIND PERPENDICULAR TO UPWIND FACE (Wilson, 1979).


Plume boundary
Isopleth of maximum allowable ambient concentration
x < x_m
A
Plume centerline
x > x_m
Δh
U_S
Windspeed at stack height
h_S
Stack
Z_M
x
X_A

(a) Definition of minimum descent height Z_m.



(b) Typical minimum descent height calculation for one stability category

FIGURE 3. MINIMUM DESCENT HEIGHT OF MAXIMUM ALLOWABLE CONCENTRATION (Wilson & Netterville, 1976).

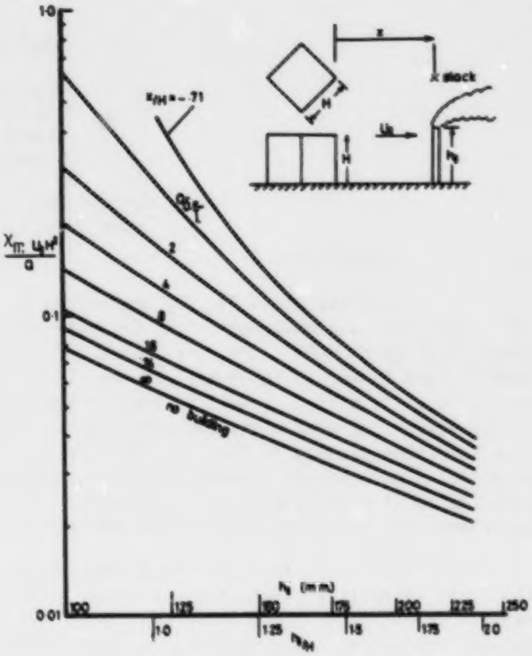
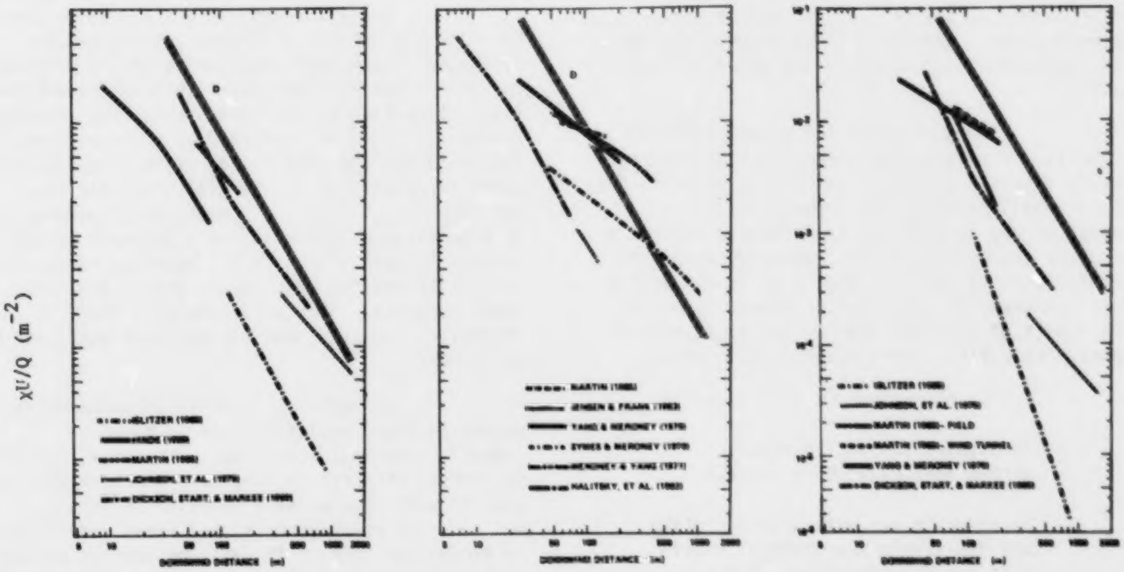


FIGURE 4. EFFECT OF BUILDING/STACK SEPARATION ON MAX. GROUND-LEVEL CONCENTRATION FOR VARYING STACK HEIGHTS DOWNWIND OF BUILDING AT ANGLE TO WIND (Barrett et al., 1978).



(a) Field tests, neutral to unstable conditions

(b) Wind tunnel tests, neutral & unstable conditions

(c) Field & wind tunnel tests in stable conditions

FIGURE 5. FAR WAKE CONCENTRATION DECAY BEHIND BUILDING CLUSTERS. HATCHED LINES ARE P-G "D" AND "F" ESTIMATES FOR COMPARISON (Abbey, 1976).

Appendix H

MODELING VISIBILITY

Douglas A. Latimer
Robert W. Bergstrom
Clark D. Johnson
James P. Killus

Systems Applications, Incorporated
San Rafael, California 94903

1. INTRODUCTION

With the Clean Air Act Amendments of 1977, Congress gave government [the Environmental Protection Agency (EPA), the states, and the Federal Land Managers (FLM)] the responsibility to protect the visibility as well as the air quality of the scenic national parks and wilderness areas that have been designated Class I. The EPA is expected to propose and promulgate regulations in 1980 that will provide a mechanism by which government can carry out these responsibilities.

Currently it is not clear what form the EPA's visibility regulations will take. For example, regulations could establish limits on ambient concentrations of particulate matter and nitrogen dioxide (NO_2), which are responsible for the visual impacts, or they could place limits on visibility-related optical parameters such as extinction coefficient, visual range, contrast, blue-red ratio and ΔE (Latimer et al., 1978). Alternatively, there is some evidence (Latimer, 1979), that existing New Source Performance Standards (NSPS) and Prevention of Significant Deterioration (PSD) legislation may be adequate to protect Class I area visibility without the need for regulations specifically addressed to visibility.

If regulations are promulgated that place limits on physical parameters that characterize visibility, then a visibility modeling capability will be needed to predict the visual impact of new sources and to determine whether a proposed source will be in compliance with such regulations. Visibility modeling, in conjunction with measurements, may also be needed to assess the impact of existing sources and to determine whether retrofit control technology is needed.

The objectives of this paper are

- > To discuss visibility modeling within this regulatory context.
- > To describe a simple SO_2-to-sulfate and NO_2-to-nitrate chemical model that has been developed recently.
- > To illustrate the use of visibility models for evaluating impacts of new, or existing, sources.
- > To discuss the current uncertainties and needs for additional work.

2. MODELING WITHIN THE REGULATORY CONTEXT

Figure 1 illustrates schematically the regulatory actions necessary to restore and protect visibility in Class I areas. The essential feedback or control function is provided by those government organizations responsible for visibility protection. This function is shown in Figure 1 as the diamond. From a set of regulations and criteria, a government official must decide whether the visibility impairment from a project, as predicted by a model or measured by the human observer or with scientific instruments, is acceptable. If it is, permits are granted for the project; if not, emissions controls, alternative siting, or project cancellation must be considered.

Model assessments are the only means by which the impacts of future actions can be evaluated. Each model assessment should provide the government official with sufficient quantitative information to indicate whether the proposed action will have an "adverse" or a "significant" impact (using the terminology of Sections 165 and 169A, respectively, of the 1977 Clean Air Act Amendments). A final determination of adversity or significance then requires a judgment of the balance between the aesthetic benefits of good visibility and the costs necessary to achieve those benefits. Such a cost-benefit analysis for visibility, however, has not yet been performed by government.

In addition, a criteria document is needed so that regulations can be based on specific numerical standards. Otherwise, how is a government official to determine the adversity or significance of a model prediction of impact, defined, for example, by visual range reduction of 10 percent or more on 15 days per year or perceptible plumes having ΔE values greater than 5 on 7

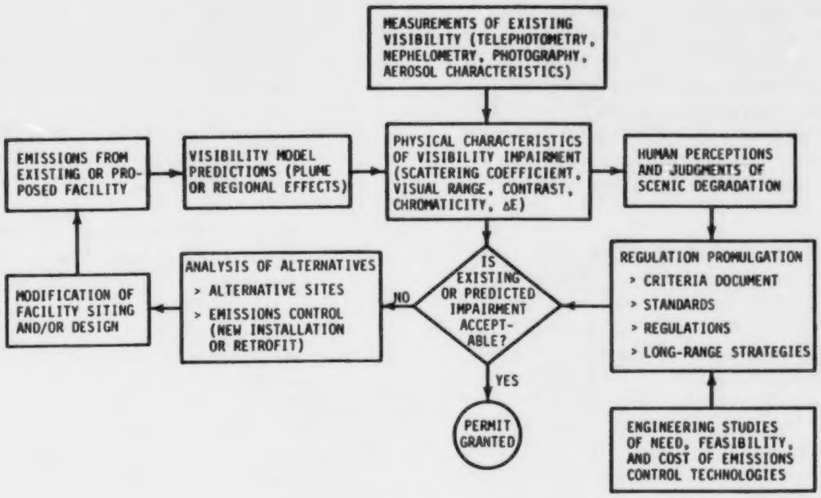

MEASUREMENTS OF EXISTING VISIBILITY (TELEPHOTOMETRY, NEPHELOMETRY, PHOTOGRAPHY, AEROSOL CHARACTERISTICS)
EMISSIONS FROM EXISTING OR PROPOSED FACILITY
VISIBILITY MODEL PREDICTIONS (PLUME OR REGIONAL EFFECTS)
PHYSICAL CHARACTERISTICS OF VISIBILITY IMPAIRMENT (SCATTERING COEFFICIENT, VISUAL RANGE, CONTRAST, CHROMATICITY, ΔE)
HUMAN PERCEPTIONS AND JUDGMENTS OF SCENIC DEGRADATION
MODIFICATION OF FACILITY SITING AND/OR DESIGN
ANALYSIS OF ALTERNATIVES
> ALTERNATIVE SITES
> EMISSIONS CONTROL (NEW INSTALLATION OR RETROFIT)
NO
IS EXISTING OR PREDICTED IMPAIRMENT ACCEPTABLE?
YES
PERMIT GRANTED
REGULATION PROMULGATION
> CRITERIA DOCUMENT
> STANDARDS
> REGULATIONS
> LONG-RANGE STRATEGIES
ENGINEERING STUDIES OF NEED, FEASIBILITY, AND COST OF EMISSIONS CONTROL TECHNOLOGIES

Figure 1. Schematic logic diagram of regulatory actions necessary to restore and protect visibility in Class I areas

days per year? Each time a visibility model assessment is reviewed, a unique set of judgment criteria might be used. It is not difficult to imagine that the uncertainty created by regulations without clearly defined criteria and standards could cause a bureaucratic nightmare. In addition, industry would not be able to plan ahead with assurance. The modeling community must also know what parameters will be used to evaluate emissions source impacts so that models can be designed to evaluate the magnitude and frequency of occurrence of visibility impairment with the appropriate measures and so that model assessments can be used to determine compliance with visibility regulations. It is hoped that the impending EPA regulations will address these areas of uncertainty.

We believe that regulations could be devised to protect visibility values in Class I areas without specifically regulating optical parameters descriptive of visibility (Latimer, 1979). Existing NSPS and PSD regulations by themselves may be adequate protection of visibility. Alternatively, these standards might be supplemented with Class I area plume centerline particulate and NO_2 concentration increments, which are selected to protect visibility values.

If, however, regulations are developed to limit the values of optical parameters, we recommend that two different types of visibility standards be established. The first would set limits on the increased haziness in a Class I area caused by increases in fine particulate matter. Such a standard could be written in terms of extinction coefficient, visual range, or contrast, all of which are mathematically interrelated. The second type of standard would set limits on the degree of plume discoloration visible from within a Class I area. This type of standard should be written in terms of parameters that characterize color differences of plumes relative to their viewing backgrounds such as blue-red ratio, contrast, and ΔE.

Whether visibility regulations are written in terms of ambient concentrations of pollutants or optical parameters such as contrast, criteria and standards should be based on detailed studies of the costs needed to achieve a given aesthetic benefit. Also, screening techniques should be developed so that visibility model assessments of impacts do not have to be performed for every source. For example, visiblity impact assessments should not be required for small emissions sources, sources located at distances from a Class I area approaching the mean visual range for that area, or sources isolated from Class I areas. by significant topographic features.

3. THE STRUCTURE OF VISIBILITY MODELS

Several visibility models are currently available for use (Latimer and Samuelsen, 1975, 1978; Latimer et al., 1978; Williams, Wecksung, and Leonard, 1978; Ozkaynak, Isaacs, and Murphy, 1979; Sorge and Mohnen, 1979; Norberg, Husar, and Patterson, 1978; Bowers, 1979). Many air quality modelers have developed their own visibility models, using atmospheric optics theory coupled with some type of air quality simulation model, either a plume or a regional model. All of these models are somewhat similar in structure.

Under contract with the EPA, Systems Applications, Incorporated (SAI) in 1978 developed both plume and regional models to predict visibility impairment caused by air pollution (Latimer et al., 1978). SAI's plume visibility model is currently being refined, evaluated by comparison with field measurements, and documented for general use, again under contract with the EPA. This refined model will be available for use later in 1980. Current plans are that this model, a user's guide to the model, and other technical guidance documents, including a document entitled "Workbook for Estimating Visibility Impairment," similar to the Turner (1970) workbook, will be issued by the EPA when regulations are promulgated.

The main features of SAI's visibility models are briefly summarized below. Readers who are interested in the technical details of the models should refer to SAI's report to the EPA, "The Development of Mathematical Models for the Prediction of Anthropogenic Visibility Impairment," (Latimer et al., 1978). Some of the recently developed features of SAI's plume visibility model are described in the next section of this paper, followed by a description of an application of the model and an assessment case study.

Figure 2 is a schematic flow diagram of the main elements of visibility models such as SAI's models. Modeling of visibility impairment requires mathematical descriptions for the following physical and chemical atmospheric processes in succession:

> Emissions.

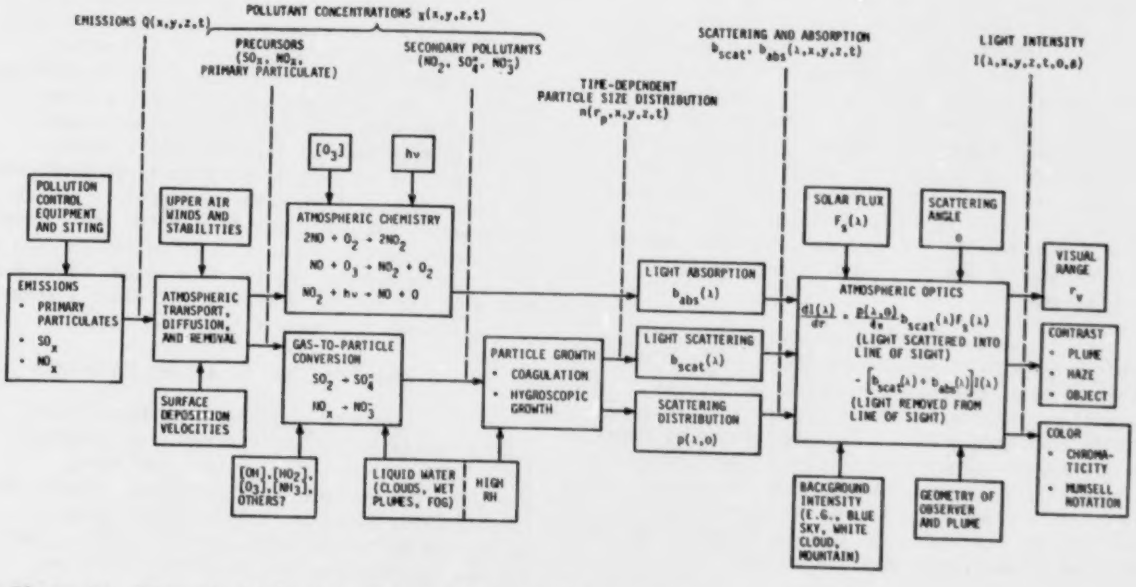
> Pollutant transport, diffusion, and removal.

> Chemical reactions and physical transformations of precursors in the atmosphere.

> Light scattering and absorption properties of aerosols and gases resulting from precursors.

> Radiative transfer through aerosols and gases along different lines of sight.

Of the two different types of models that SAI has developed, the Gaussian plume model, which is used in this study, is designed to calculate the visual impact of emissions from a single point source at downwind distances on the order of 100 km. The regional grid model is designed to calculate the visual impact resulting from multiple sources over several days within a 1000 km x 1000 km region.



EMISSIONS $Q(x,y,z,t)$
POLLUTANT CONCENTRATIONS $\chi(x,y,z,t)$
PRECURSORS (SO_x, NO_x, PRIMARY PARTICULATE)
SECONDARY POLLUTANTS (NO_2, $SO_4^=$, NO_3^-)
TIME-DEPENDENT PARTICLE SIZE DISTRIBUTION $n(r_p,x,y,z,t)$
SCATTERING AND ABSORPTION b_{scat}, $b_{abs}(\lambda,x,y,z,t)$
LIGHT INTENSITY $I(\lambda,x,y,z,t,\theta,\phi)$
POLLUTION CONTROL EQUIPMENT AND SITING
EMISSIONS
• PRIMARY PARTICULATES
• SO_x
• NO_x
UPPER AIR WINDS AND STABILITIES
ATMOSPHERIC TRANSPORT, DIFFUSION, AND REMOVAL
SURFACE DEPOSITION VELOCITIES
$[O_3]$
$h\nu$
ATMOSPHERIC CHEMISTRY
$2NO + O_2 \rightarrow 2NO_2$
$NO + O_3 \rightarrow NO_2 + O_2$
$NO_2 + h\nu \rightarrow NO + O$
GAS-TO-PARTICLE CONVERSION
$SO_2 \rightarrow SO_4^=$
$NO_x \rightarrow NO_3^-$
$[OH]$, $[HO_2]$, $[O_3]$, $[NH_3]$, OTHERS?
LIQUID WATER (CLOUDS, WET PLUMES, FOG)
HIGH RH
PARTICLE GROWTH
• COAGULATION
• HYGROSCOPIC GROWTH
LIGHT ABSORPTION $b_{abs}(\lambda)$
LIGHT SCATTERING $b_{scat}(\lambda)$
SCATTERING DISTRIBUTION $p(\lambda,\theta)$
SOLAR FLUX $F_s(\lambda)$
SCATTERING ANGLE θ
ATMOSPHERIC OPTICS
$\frac{dI(\lambda)}{dr} = \frac{p(\lambda,\theta)}{4\pi} b_{scat}(\lambda) F_s(\lambda)$
(LIGHT SCATTERED INTO LINE OF SIGHT)
$- \left[b_{scat}(\lambda) + b_{abs}(\lambda)\right] I(\lambda)$
(LIGHT REMOVED FROM LINE OF SIGHT)
BACKGROUND INTENSITY (E.G., BLUE SKY, WHITE CLOUD, MOUNTAIN)
GEOMETRY OF OBSERVER AND PLUME
VISUAL RANGE r_v
CONTRAST
• PLUME
• HAZE
• OBJECT
COLOR
• CHROMATICITY
• MUNSELL NOTATION

Figure 2. Schematic logic flow diagram of the visibility models

Distinctly different problems are handled by plume and regional visibility models. At distances as far as approximately 100 km from a point source such as a power plant, a plume itself may be perceptible because it is colored differently or is darker or brighter than the background. In such cases, the visual impact is not the result of a reduction in visual range (though this may occur), but rather of atmospheric discoloration in a well-defined plume. Beyond about 100 km downwind, however, individual plumes generally become uniformly dispersed in the vertical direction within the mixed layer and are no longer distinguishable as plumes. On a regional scale, visual impact is caused by a general haziness due to reduced visual range.

It appears that the first problem, impacts caused by plumes from single, identifiable sources, will be the only concern addressed with the first phase of visibility regulations; regional haze will not be addressed in the first-phase regulations (EPA, 1979; Federal Register, 1979). Accordingly, the discussion in the following sections of this paper will concern the plume visibility model, designed to assess the visual impacts of plumes from a single emissions source such as a power plant.

4. GAS-TO-PARTICLE CONVERSION IN PLUMES

For most models to date, simplifying assumptions are made regarding gas-to-particle conversion (Figure 2 shows the role of this process in visibility models). The usual approach is to assume that sulfur dioxide (SO_2) gas is converted to sulfate ($SO_4^=$) aerosol at some constant rate; this approach employs a user-input value of a pseudo-first-order rate constant whose value is empirically determined. There is considerable variation in such rates, however, with measured SO_2-to-$SO_4^=$ conversion rates ranging from a few tenths of a percent to several percent per hour.

Much of the variance in SO_2-to-$SO_4^=$ conversion observed in field measurement programs in recent years can be explained using a model that accounts for the reactions of plume SO_2 and NO_2 with the hydroxyl (OH•) radical. In clean background areas the gas-phase oxidation of SO_2 and NO_2 to sulfate aerosol and nitrate (nitric acid vapor) is due primarily to the reaction of these species with OH•. Although previous assessments of homogeneous (gas-phase) oxidation of SO_2 to sulfate estimated the proportion assignable to the reaction with hydroxyl to be from about 75 percent in clean atmospheres (Calvert, et al. 1978; Altshuller, 1979) to as low as 40 percent in polluted urban air (Isaksen, Hesstredt, and Hov, 1978), more recent estimates place these values much higher. The kinetic models forming the basis of the early estimates used the value of $1.3 \times 10^{-12} cm^3 mol^{-1} s^{-1}$ for the rate constant of reaction for HO_2 and CH_3O_2 with NO. More recently, however, this rate has been measured at $8.1 \times 10^{-12} cm^3 mol^{-1} s^{-1}$ (Hampson and Garvin, 1978). This larger rate constant lowers the expected concentration of these peroxy radicals by a factor of 6 and, in turn, greatly reduces the SO_2 conversion resulting from reactions with these radicals. When recalculated using the new rate constant, the fraction of SO_2-to-sulfate conversion that results from reaction with the hydroxyl radical is approximately 95 percent for clean atmospheres and 70 percent for the extremely polluted case.

These estimates are supported by the work of Miller (1978), who found that the SO_2 oxidation rate was not dependent on the absolute concentrations of hydrocarbons and nitrogen oxides but on the ratio of nonmethane hydrocarbon to nitrogen oxides.

The rate of sulfate (and nitric acid) formation can be estimated by calculating the steady-state concentration of OH• within a plume. This steady-state plume OH• concentration is calculated by balancing the rate of OH• production with the rate of OH• destruction. The following reactions are used:

> OH• production.*

- $O_3 + h\nu \rightarrow O('D) + O_2$ (K_1 = function of UV radiation) . (1)
- $O('D) + M \rightarrow O(^3P) + M$ ($K_2 = 4.45 \times 10^{10} min^{-1}$) . (2)
- $O('D) + H_2O \rightarrow 2\ OH\cdot$ ($K_3 = 3.4 \times 10^5 ppm^{-1} min^{-1}$) . (3)

> OH• destruction (major sinks in plumes).

- $OH\cdot + SO_2 \rightarrow HSO_3 \ldots$ ($K_4 = 2.0 \times 10^3 ppm^{-1} min^{-1}$) . (4)
- $OH\cdot + NO_2 \rightarrow HNO_3$ ($K_5 = 1.4 \times 10^4 ppm^{-1} min^{-1}$. (5)

With the assumption of steady-state concentrations of O('D) and OH• in the plume, we can write the following equations:

$$\frac{d[OH\cdot]}{dt} = \frac{d[O'(D)]}{dt} = 0 \quad , \tag{6}$$

$$\frac{d[O('D)]}{dt} = K_1(h\nu)[O_3] - K_2[O('D)] - K_3[O'(D)][H_2O] \ , \tag{7}$$

$$[O('D)] = \frac{K_1(h\nu)[O_3]}{K_2 + K_3[H_2O]} \quad , \tag{8}$$

$$\frac{d[OH\cdot]}{dt} = 2\ K_3[O('D)][H_2O] - K_4[OH\cdot][SO_2] - K_5[OH\cdot][NO_2] \ . \tag{9}$$

$$[OH\cdot] = \frac{2\ K_3[O('D)][H_2O]}{K_4[SO_2] + K_5[NO_2]} \quad . \tag{10}$$

With this steady-state concentration, plume pseudo-first-order SO_2-to-SO_4 and NO_2-to-HNO_3 conversion rates can be calculated as follows:

---

* See the Appendix.

$$-\frac{1}{[SO_4]}\frac{d[SO_2]}{dt} = K_4[OH\cdot] \quad , \qquad (11)$$

$$-\frac{1}{[NO_2]}\frac{d[NO_2]}{dt} = K_5[OH\cdot] \quad . \qquad (12)$$

Note from Eqs. (8) and (10) that power plant plume OH• concentrations are reduced below background tropospheric values for two reasons:

- Plume ozone (O_3) concentrations are reduced below background values because of the reaction $NO + O_3 \rightarrow NO_2 + O_2$ (Latimer et al., 1978).
- Plume concentrations of NO_2 and SO_2 are high, thus reducing steady-state OH• concentrations.

It should be pointed out that at night there is no production of OH• from ozone photolysis; also, in early morning and late afternoon and in winter OH• production is diminished because ultraviolet flux decreases as solar zenith angles approach 90°. Thus, sulfate and nitrate are not formed at night and are formed only very slowly in concentrated plumes.

The visibility model was run, with this OH• chemical mechanism included, for a plume from a 1600 Mwe coal-fired power plant. On the basis of this case study, the effect of the following on sulfate formation rate was studied:

- Plume dilution, corresponding to neutral and isothermal stable conditions [using TVA σ_y, σ_z values (Carpenter et al., 1971) at various downwind distances.
- Time of day.
- Season.

The history of sulfate formation in a plume parcel based on this case study is shown in Figure 3. Note that time of day, season, and plume concentration all affect the rate of sulfate formation. This effect is explained by the formulations shown in Eqs. (8), (10), and (11).* The SO_2-to-$SO_4^=$ conversion rate is directly proportional to the plume ozone and water vapor concentrations and to ultraviolet (UV) flux, and it is inversely proportional to plume SO_2 and NO_2 concentrations.

The effect of UV flux is noted in both the seasonal and time-of-day dependence of the sulfate formation rate. Note from Figure 3 that no conversion of SO_2 to sulfate occurs at night. Sulfate formation starts after sunrise, reaches its maximum rate at noon (when ultraviolet flux reaches a maximum), and then begins to decrease. Thus, a plume parcel observed at a given distance downwind from a power plant may have undergone sulfate formation during only a fraction of its journey. For example, a plume parcel 160 km downwind of its source, when winds are 2 m/s, would require a transit time of more than 22 hours. During this period sulfate would be formed only during the sunlight hours [see Figure 3(b)].

The effect of plume SO_2 and NO_2 concentrations on the steady-state OH• concentration and hence on sulfate formation is also shown in Figure 3. Note, by comparing Figure 3(a) with 3(b) that sulfate formation is much more rapid when plumes parcels are more dilute (at farther downwind distances). In the dilute, neutral-stability plumes at downwind distances greater than 100 km, sulfate formation rates reach noon peaks of about .06, .02, and .004 percent per hour, respectively, for the summer, fall, and winter simulations for this latitude. In the less dilute, neutral-stability plumes at downwind distances of less than 100 km, the sulfate formation rates peak at .016, .008, and less than .002 percent per hour for summer, fall, and winter, respectively. However, formation rates in stable plumes are considerably less than .01 percent per hour even at noon on a summer day.

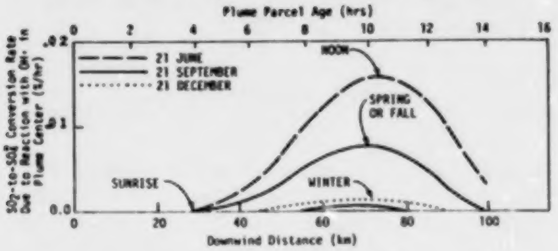
These model calculations support the argument made by Latimer (1980) that the atmospheric discoloration of stable power plant plumes (where particulate emissions are controlled) is caused primarily by nitrogen dioxide (NO_2) gas, not by secondary sulfate aerosol. Sulfate aerosol is not formed in plumes during nighttime transport, and formation is minimal during daytime transport when plumes are concentrated (as they would be during stable conditions). After plumes have become well mixed, significant rates of sulfate formation can occur in the daytime, although these rates approach zero with low sun angles. Thus, sulfate formation is a long range (greater than 100 km) multiday phenomenon, not a near-source problem.

5. CHARACTERIZING THE FREQUENCY DISTRIBUTION OF PLUME VISIBILITY IMPAIRMENT

As discussed in Section 2 of this paper, the first regulatory applications of visibility models will be the assessments of the impacts of single, existing or proposed sources on visibility in Class I areas, not the impacts of multiple sources on regional haze. For such a visual impact assessment, it is important to determine the frequency of occurrence of visual impact, because the adversity or significance of

---

* It should be noted that sulfate can form not only by reactions with OH• but also on existing, catalytic fly ash particles. However, this sulfate coating on existing particles is not expected to significantly add to aerosol surface area, which is important for light scattering.



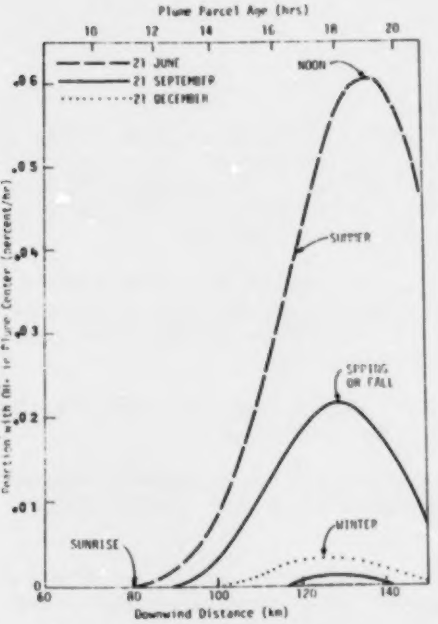
Notes:

(1) Time-dependent SO_2-to-$SO_4^=$ conversion rates for the history of a parcel in the center of a well-mixed plume (neutral stability) which is observed 100 km downwind at 1500 (in the afternoon) are plotted for indicated seasons in Figure 3(a).

(2) Comparable conversion rates for a well-mixed plume observed 960 km downwind at 1500 are plotted in Figure 3(b).

(3) Conversion rates for stable plumes (isothermal stability) are considerably less than 0.1 percent/hr. See shaded areas at the bottom of graphs.

(a) Life history of plume parcel observed 100 km downwind at 1500 hours



(b) Life history of a plume parcel observed 160 km downwind at 1500 hours

Figure 3. Calculated time dependence of sulfate formation in the center of plumes from a 1600 MWe coal-fired power plant, 2 m/sec winds, neutral and isothermal stability conditions

impact is dependent on how frequent the impact of a given magnitude occurs. For example, if a plume were perceptible from a Class I area a third of the time the impact would be considered much more significant than if it occurs only one day per year. The determination of frequency of occurrence of impact should be an integral part of a visibility impact assessment.

In this section we discuss how the models described earlier can be applied to determine both the magnitude and frequency of occurrence of visual impact.* This procedure entails making several model runs for different values of the following important input parameters:

> Emissions rates (particulate, SO_2, NO_x).

> Wind speed.

> Wind direction.

> Atmospheric stability.

> Mixing depth.

> Background ozone concentration.

> Background visual range.

> Ultraviolet solar intensity.

> Time of day and season.

> Orientation of observer, plume, and sun.

> Viewing background (whether it is sky, cloud, or snow-covered, sunlit, or shaded terrain).

Because of the large number of variables important to a visual impact calculation, the visibility computer model must be applied several times to assess the magnitude and frequency of occurrence of visual impact. It would be ideal to calculate hourly impacts over the course of a year or more using hourly values of the above variables. However, such an extensive data base is rarely available for use. Even if it were available, the computing costs involved would be prohibitive. It is therefore preferable to select a few representative, discrete values of each of these variables in order to represent the range (i.e., the magnitude and frequency of occurrence) of visual impact over a given period of time, such as a season or year. One can start with the conditions that cause the worst impacts and then assess the frequency of occurrence, in a season or year, of all the variables having worst-case values simultaneously. It is possible to imagine a worst-case impact condition that would never occur in the real atmosphere; this condition could be represented on a cumulative frequency plot, such as that of Figure 4, as point A. The impact is great, but it virtually never occurs. If another worst-case situation less extreme than point A were selected, the magnitude of impact would be less, but it might occur with some nonzero frequency. It is possible to select various values of all the important input variables and to assess the frequency with which conditions, resulting in impacts worse than a given impact, would occur. By this process, several points necessary to specify the frequency distribution could be obtained: for example, points B, C, and D in Figure 4. With average (50th percentile) conditions, a negligible impact, as shown at point E in Figure 4, might be found. In Figure 4, the ordinate could be any of the parameters used to characterize visibility impairment, such as visual range reduction, plume contrast, blue-red ratio, or ΔE, and the abscissa could represent cumulative frequency over a season or a year.

---

* Plume visual impact is considered to be the result of visual range reduction (which results in the reduction of contrast between landscape features), and plume appearance (which results from contrast and color differences between the plume and the viewing background).

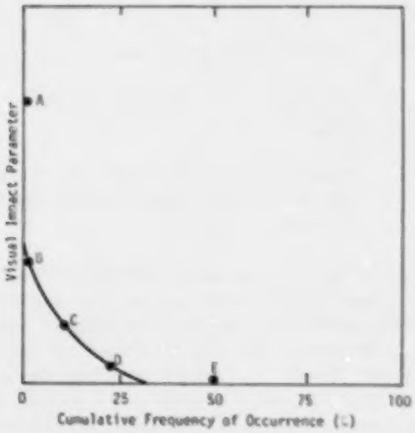


Figure 4. Example of a frequency distribution of visual impact

In a visual impact assessment it is recommended that one select various permutations of upper air wind speed, wind direction, and atmospheric stability; background ozone concentration; and background visual range to specify the frequency distribution of visual impact. If one has a large concurrent data base of all five of these variables, it would be desirable to calculate a five-way joint probability distribution matrix and to use these joint probabilities to calculate frequency of occurrence of impact. However, in most situations such a data base is not available, and one must treat the various worst-case events as independent probabilities. With this assumption, the probability of worst-case impacts can be calculated by multiplying the independent probabilities. This can be represented as follows:

$$f(y > y') = \prod_i f(x_i > x_i') \quad ,$$

where $f(y > y')$ is the cumulative frequency of impact y great than y', and $f(x_i > x_i')$ is the cumulative frequency of variable x_i having values that would cause greater impact than the value x_i'.

In such an application, one might obtain an estimate of cumulative frequency by using the joint frequency distribution of upper air wind speed and wind direction and the separate frequency distributions of upper air stability, ozone concentration, and visual range. For example,

$$f(\Delta E > \Delta E') = f(u < u', WD < WD') \cdot f(s > s') \cdot f([O_3] > [O_3]') \cdot f(r_v > r_v')$$

where

$f(\Delta E > \Delta E')$ = the frequency of occurrence of ΔE values greater than ΔE'. ΔE' is calculated on the basis of a wind speed u', wind direction WD', stability s', ozone concentration $[O_3']$, and visual range r_v'.

$f(u < u', WD < WD')$ = the frequency of occurrence of wind speeds less than u' associated within a specified value (WD') of the worst-case wind direction.

$f(s > s')$ = the frequency of occurrence of stabilities greater than s'.

$f([O_3] > [O_3]')$ = the frequency of occurrence of background ozone concentrations greater than $[O_3]'$.

$f(r_v > r_v')$ = the frequency of occurrence of background visual range values greater than r_v'.

Each of the input parameters that are important to the visibility model calculation vary significantly over a period of a year and are discussed in the following paragraphs.

5.1 Wind Speed

Wind speed affects plume visual impact strongly because plume centerline concentrations and plume line-of-sight integrals are inversely proportional to wind speed. Greater impact would be expected during light-wind stagnation conditions than during strong-wind, well-ventilated conditions. Also, since the age of a plume parcel at a given distance downwind from a power plant is inversely proportional to wind speed, more time is available at low wind speeds for the chemical conversion of primary emissions. A well-aged plume parcel is more likely to cause a reduction in visual range than is a younger one. However, the time necessary to transport emissions a given distance toward a Class I area increases with decreasing wind speed. Thus, during light-wind conditions several hours persistence may be needed to transport emissions to a Class I area where they could cause visual impact.

5.2 Wind Direction

Wind direction also affects plume visual impact because the direction of plume parcel transport affects the orientation of the plume with respect to the observer. If the plume is transported directly toward an observer, the observer's lines of sight directly along the plume center will be significantly affected. If the observer's line of sight is oblique to or along the plume axis, visual impact will be greater than if the line of sight is normal to the plume axis. Also, the direction of plume transport affects the distance between the observer and the plume material. Plume discoloration is diminished by light scattered by the intervening, or background, atmosphere. The more distant the plume material, the less colored and less perceptible it is likely to be. This decrease in plume coloration can be expressed approximately as follows:

$$\Delta E(r) = \Delta E(0) \exp(-Kr/r_{v_0}) \quad , \qquad (13)$$

where

$\Delta E(r)$, $\Delta E(0)$ = plume perceptibility due to color differences at plume-observer distances r and 0, respectively.

K = a constant.

r = plume-observer distance.

r_{v_0} = background visual range.

It should be noted that visual range reduction will not decrease with increasing distance between the plume and the observer. However, the aesthetic effects of this visual range reduction would be expected to be smaller as the plume-observer distance increases, because only distant terrain objects would be affected and fewer lines of sight would be impacted. Also, the objectionability of plume discoloration is likely to be less if the plume is farther away, because the plume will appear smaller (i.e., fewer lines of sight will be affected).

To illustrate the effect of wind direction, Figure 5 shows the positions of two plumes from a hypothetical power plant relative to a vista in a Class I area 90 km away from the power plant. Plume 2-σ_y outlines are shown to scale with σ_y values appropriate for a TVA neutral stability (Carpenter et al., 1971). Of course, actual plume trajectories will be affected by wind channeling, complex terrain, and changes in wind direction with time, so these figures are only idealized representations. Figure 5 shows trajectories that could occur with north-northwesterly and north-northeasterly winds. A plume associated with a north-northeasterly wind direction (defined by a sector 22.5° wide) could be anywhere within the extremes of the sector shown in Figure 5. Thus, a wide range of impacts could occur, associated with a north-northeasterly winds. The worst case would be that shown in Figure 5(a), in which the plume is transported directly toward the observer. Of course, the worst-case conditions of Figure 5(a) would occur for only some of the periods of north-northeasterly flow. Figure 5(b) shows that, for another wind direction, plume discoloration would be considerably less because plumes would be tens of kilometers away from the observer, and the observer's line of sight could be nearly perpendicular to the plume, not along the plume as in Figure 6(a).

5.3 Atmospheric Stability

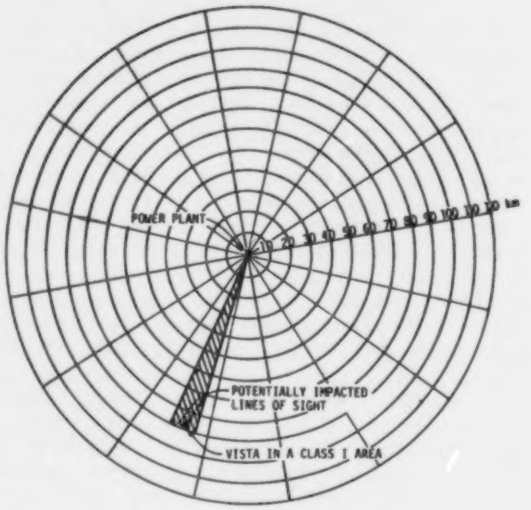
Upper air stability controls the rate at which power plant emissions are mixed with ambient air. During stable conditions, diffusion is limited, particularly in the vertical direction, so plumes remain as ribbon-like layers. Plume discoloration is most apparent during such stable conditions because the integral of NO_2 and particulate concentrations along the line of sight is greater. During well-mixed (neutral or unstable) conditions, plumes are rapidly dispersed and not likely to be visible as plumes per se.

Stability, or the rate of plume mixing, also has an affect on chemical conversion within a plume. The conversion of nitric oxide (NO) to nitrogen dioxide (NO_2) is diffusion-limited in stable plumes as is sulfate and nitrate formation, as was discussed in the previous section of this paper.

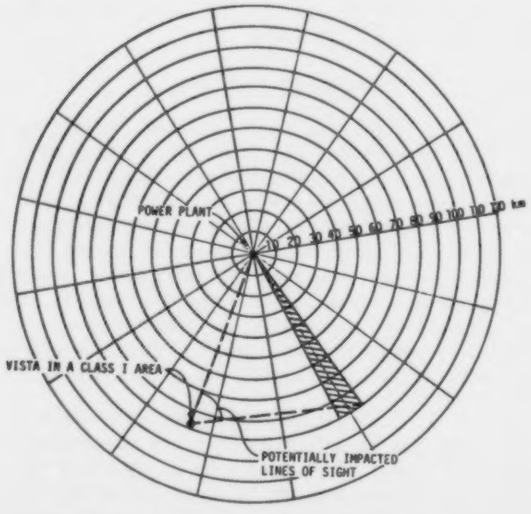
5.4 Background Ozone Concentration

An important input parameter to the visibility model is the background ozone concentration, that is, the concentration of ozone outside the plume. Ozone reacts directly with the colorless nitric oxide emitted from power plants to form the brownish gas nitrogen dioxide, the principal plume colorant:

$$NO + O_3 + NO_2 + O_2$$


POWER PLANT
10 20 30 40 50 60 70 80 90 100 110 120 km
POTENTIALLY IMPACTED LINES OF SIGHT
VISTA IN A CLASS I AREA

(a) Worst-case wind direction


POWER PLANT
10 20 30 40 50 60 70 80 90 100 110 120 km
VISTA IN A CLASS I AREA
POTENTIALLY IMPACTED LINES OF SIGHT

(b) Wind direction resulting in less impact than the worst case

Figure 5. Schematic diagram showing plume-observer geometries for two different wind directions

Ozone is also important indirectly in the oxidation of plume NO_2 and SO_2 since ultraviolet radiation photolyzes ozone to form the hydroxyl radical (OH•) that reacts with NO_2 and SO_2 to form nitric acid and sulfate aerosol, as discussed previously.

Since the background ozone concentrations at plume altitude, often several hundred meters above ground, is needed for model input, one must interpret ground-level ozone concentration data with care. In their analysis of long term ozone concentration data at remote U.S. sites, Singh, Ludwig, and Johnson (1978) reported that there is a significant diurnal variation in ozone concentrations at the surface because of the surface depletion of ozone. They reported a significant reservoir of ozone in the free troposphere varying in concentration from about 30 ppb in the winter to about 60 ppb in the summer. Tropospheric ozone is rapidly mixed to the ground during the daytime; this causes surface concentrations near the free troposheric value. However, at night and in early morning, ozone is no longer mixed to the ground because of the development of a ground-based stable layer. During this period, ground-level ozone concentrations gradually decrease as a result of a surface depletion mechanism. In relatively remote unpolluted regions, one would not expect a significant anthropogenic source of ozone. Rather, surface ozone appears to be affected by downward mixing from the free troposphere.

Since the ozone concentrations at plume altitude are of concern in visibility calculations, it is appropriate to use the daily maximum value of the surface concentration to represent the daily average concentration at plume altitude. Calculations should be made for the median (50th percentile) ozone concentration as a minimum and possibly for the 25th and 75th percentiles also.

5.5 Background Visual Range

Background visual range is also an important input parameter because the magnitude of plume discoloration visible from a given location depends on the clarity of the intervening atmosphere. Plume discoloration is much more noticeable in the extremely clear areas of the Southwest, for example, than in hazy areas. Equation (13) shows that as the background visual range (r_{V0}) decreases (i.e., the atmosphere becomes hazier), the degree of plume discoloration (characterized by the ΔE value) decreases also. Thus, one must supply one or more (e.g., 25th, 50th, and 75th percentiles) values of background visual range in the study area to characterize impacts for different atmospheric clarity.

5.6 Study Area Topography

The topography of an area also has an influence on visibility impairment. High terrain affects the transport of emissions, particularly during worst-case stable conditions. It is likely that a stable plume will be channeled by high terrain and remain in a valley. Thus, the assumption of a straight plume trajectory approaching an observer location on elevated terrain, such as shown in Figure 5(a), may never occur in some areas. The topography also affects the rate of dilution of plumes, with mechanically induced turbulence enhancing plume dilution.

Topography also affects the views from a given vista location. For example, topography can obstruct views in certain directions from a vista where plume material is located. It can also have an effect on the type of viewing background (what is visible behind the plume). Plume discoloration is generally less apparent when terrain is the viewing background; plumes are generally more visible against clear sky or white backgrounds, like those provided by clouds or snow-covered terrain. In the current version of the SAI plume visibility model, the distances to terrain features that are viewing backgrounds are used as input (these are functions of the azimuth of the line of sight), and it is assumed that terrain can be white, gray, or black, depending on whether it is snow-covered, in sun, or in shadow, respectively. Model calculations are also performed assuming the plume is observed against a clear sky viewing background.

5.7 Season and Time of Day

As indicated in the previous section gas-to-particle conversion is also a function of season and time of day, with higher conversion rates at times when ultraviolet flux is most intense. Also, the sun angle is dependent on season and time of day.

5.8 Model Runs

If one used all the permutations of the important input variables, one could make hundreds of plume visibility model runs to characterize the frequency distribution of visual impact over a season or a year. For example, if 5 wind directions, 3 wind speeds, 2 stabilities, 3 background visual ranges, 3 background ozone concentrations, and 2 seasons are evaluated, one would have to make 540 runs (5 x 3 x 2 x 3 x 3 x 2). Furthermore, if 10 downwind distances were evaluated and 4 viewing background colorations were considered for each line-of-sight geometry, a total of 21,600 individual line-of-sight calculations would be needed. The computing costs alone for this many calculations would be at least $10,000. To reduce costs one can reasonably approximate the frequency distribution of impacts by using median values of background ozone concentration and visual range and evaluating impacts for sun angles corresponding to one season (e.g., spring or fall), thus reducing the total number of runs in example to 30.

6. A VISIBILITY MODELING ASSESSMENT CASE STUDY

To illustrate the application of visibility models to the assessment of the visual impact of an emissions source, this section considers the impact of a hypothetical 1600 Mwe coal-fired power plant located 90 km from a Class I area. The assumption is made for this analysis that the power plant burns low sulfur coal and is equipped with 99.7 percent efficient particulate emission controls and 85 percent efficient SO_2 scrubbers; NO_x emissions are assumed to be limited by boiler combustion control techniques to 0.65 $lb/10^6$ Btu heat input, within the EPA's New Source Performance Standards.

This case study employs a revised version of the SAI plume visibility model (Latimer et al., 1978) with the following major improvements:

> Plume impacts are calculated as a function of a line-of-sight azimuth for a specific plume-observer-sun geometry. For a given scenario, plume impacts are calculated for lines of sight through the center of the plume at various downwind distances, from user-input values of:

  - The location of the emissions source (UTM coordinates and elevation msl).
  - The location of the observer (UTM coordinates and elevation msl).
  - Wind direction (azimuth in degrees).
  - Date and time.

> Plume impacts can be calculated for lines of sight directly along the plume centerline as well as for those perpendicular and oblique to the centerline.

> Sulfate (SO_4) and nitrate (HNO_3) formation rates are calculated explicitly as a function of time and space for each plume parcel, using the steady-state OH• chemical kinetics model discussed previously.

For this case study, this visibility model was applied 48 times to fully characterize the frequency distribution of visibility impairment; the model was run for all permutations of 8 wind directions, 3 wind speeds, 2 stabilities (neutral and isothermal stability), and for a median ozone concentration of 38 ppb and a median background visual range of 185 km.

Figure 6 illustrates the calculated magnitude of visual impact for 4 of these 48 ambient conditions. Each of these computer-generated plots shows four visibility impairment parameters--percentage reduction in visual range, plume blue-red ratio, plume contrast, and the plume perceptibility $\Delta E(L^*a^*b^*)$ parameter--as a function of the direction of observer's view from a given vista location in the hypothetical Class I area. The plume-observer geometries assumed for these examples are those shown in Figure 5. Figure 6(a) and 6(b) show the impacts associated with stable and neutral plumes when the wind direction is almost directly toward the observer as shown in Figure 5(a). Figures 6(c) and 6(d) show the impacts corresponding to a less severe case, with a wind direction as shown in Figure 5(b). All of the plots were based on a 2 m/sec wind speed, a 38 ppb background ozone concentration, and a 185 km background visual range. Also, calculations were performed for four different assumed viewing background colorations: (1) a clear sky, (2) a white object (characteristic of snow-covered terrain), (3) a gray object (characteristic of terrain in sun), and (4) a black object (characteristic of terrain in shadow). Furthermore, it is assumed that the distance to the terrain viewing background is a function of the direction of view (i.e., the azimuth of the line of sight). Terrain as close as 50 km is assumed as the viewing background for a 180° field of view from south to west to north. Terrain is 150 km away for azimuths from north to northeast, and a mountain 50 km away is assumed to be visible to the northeast (45° azimuth). For lines of sight to the east and southeast the plume is assumed to be viewed against clear sky only. Sun angles are calculated for morning (0900) and afternoon (1500) for 21 September. Comparing Figure 6(a) with 6(b) and Figure 6(c) with 6(d), one can see that impacts are much greater when the plume material is concentrated during isothermal stability conditions than when material is dispersed during well-mixed, neutral conditions.

Note also that the coloration of, and distance to, the terrain viewing background has an effect on the plume coloration parameters. Plumes are relatively dark and more perceptible against white terrain or clear sky viewing backgrounds and relatively bright and less perceptible against gray or black viewing backgrounds.

Plume discoloration varies from a maximum, corresponding to ΔE values greater than 30 (very discolored), to a minimum, with ΔE values as low as 1 (imperceptible). However, these plots of the magnitude of plume visual impact must be supplemented with estimates of frequency of occurrence to obtain a complete picture of the adversity or significance of impact. If plume coloration of a given magnitude occurs several days per year, the adversity or significance of that impact is greater than if it occurs less than once per year. Thus, one must determine how many days per year the magnitudes of impact shown in Figure 6 are likely to occur. For example, the impact shown in Figure 6(a) is not likely to occur more than a fraction of a percent of the time in the case study area.

As discussed previously, the frequency of occurrence of visual impact can be calculated from the joint probability of worst-case values of the following parameters:

> Wind direction.

> Wind speed.

> Stability and mixing depth.

> Background ozone concentration.

> Background visual range.

> Realistic plume transport persistence and trajectories.

For a visibility impact assessment one would use the probabilities of each of these conditions occurring concurrently to calculate cumulative frequency plots of plume coloration and visual range reduction similar to that shown in Figure 4. To characterize plume coloration, the $\Delta E(L^*a^*b^*)$ parameter has been used. The appearance of stable plumes is best characterized by this parameter because it is proportional to the degree of atmospheric discoloration (due to changes in both brightness and color hue and saturation) that causes the visual perception of plumes. Percentage visual range reduction has been used to characterize haziness caused by well-mixed plumes, which are not visible as plumes per se but nevertheless contribute to atmospheric haze and decrease the contrast of landscape features.

These frequency plots can then be used to summarize the frequency of occurrence of visual impact in various seasons, for example, as shown in Table 1.

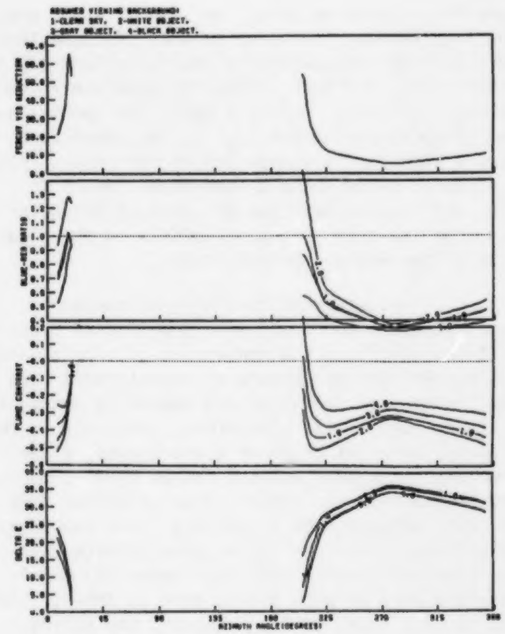
Table 1
Example of summaries of the frequency of occurrence of visual impact in a Class I area from a hypothetical power plant

(a) Perceptible plume discoloration

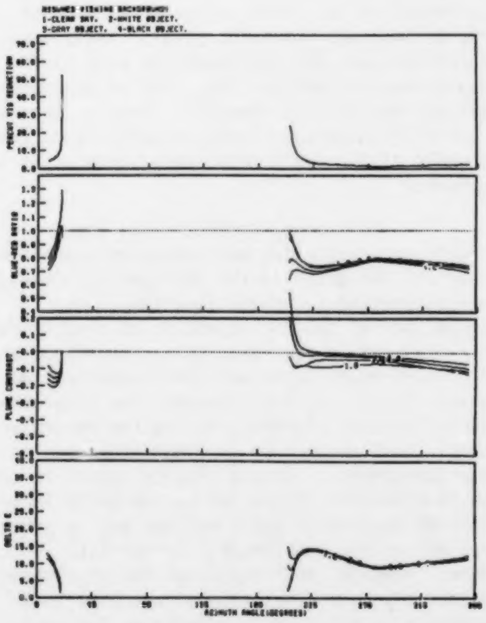
Season	Number of Mornings with $\Delta E(L^*a^*b^*)$ Greater than Indicated Value		
	2.5	5	10
Winter	6	3	1
Spring	2	1	0
Summer	3	1	0
Fall	5	2	1
Annual Total	16	7	2

(b) Haze (visual range reduction)

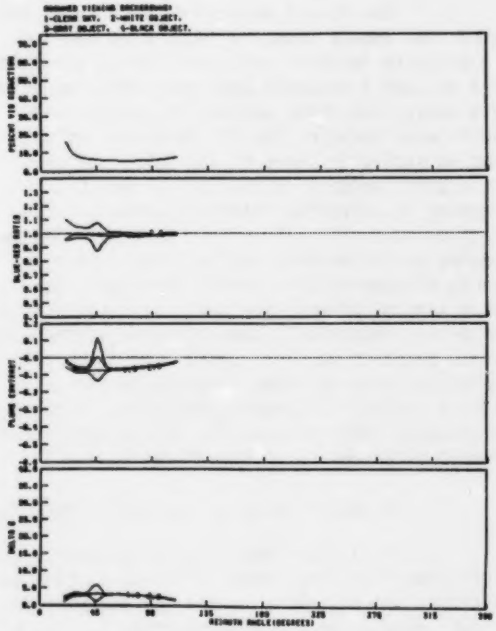
Season	Number of Days with Visual Range Reduction Greater than Indicated Value		
	5%	10%	15%
Winter	10	5	1
Spring	3	2	< 1
Summer	4	1	0
Fall	5	3	1
Annual Total	22	11	2



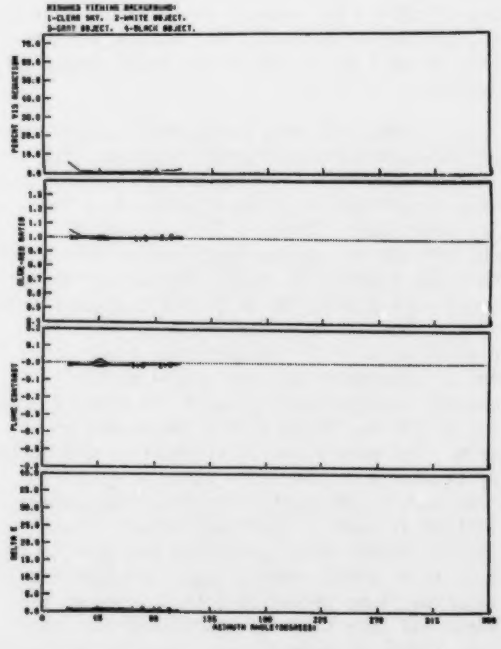
(a) North-northeast wind, isothermal stability (worst-case)



(b) North-northeast wind, neutral stability



(c) North-northwest wind, isothermal stability



(d) North-northwest wind, neutral stability

Figure 6. Plume visibility impacts for various wind directions and stabilities from the vantage of an observer in a Class I area

How does a government official judge whether the impact shown in Table 1 is adverse? If a criteria document and quantitative standards based on such a document were available, he or she would simply determine whether the calculated impacts were greater than the standard, which might be stated in terms of the number of days that a given value of a visibility impact parameter is exceeded. Without such standards, the government official is left to his or her own judgment to decide whether the given impact is adverse or significant. He or she might view photographs of plumes with measured discoloration similar to predictions, computer-generated plume-terrain perspective plots (Hayes, 1979) or color simulations such as those produced by the Los Alamos Scientific Laboratory (Williams, Wecksung, and Leonard, 1978) in order to make a subjective judgment of adversity or significance.

7. RESEARCH NEEDS IN VISIBILITY MODELING

Much additional work is needed to better understand the impact of air pollution on visibility and aesthetic values, to improve and validate models, and to develop effective, workable regulations. Much of this work is underway, but much is left to be done.

Since errors and uncertainties in each of the components of a visibility model (see Figure 2) affect the overall performance of the model, the performance of individual model components as well as of the entire model should be evaluated.

One such model component is the atmospheric transport and diffusion component (the first component shown in Figure 2). Probably the largest single source of uncertainty in a visibility model calculation is this component. Visual effects are strong functions of atmospheric dispersion; plumes from power plants, for example, are much more discolored in stable transport conditions. Thus, it is important to predict the rate of plume dilution accurately. Also, it is essential to predict mesoscale transport accurately because visual impacts in Class I areas as far as 100 to 200 km from a source may be of concern. The prediction of transport and dilution is particularly difficult in cases with complex terrain such as the western United States where visibility is such an important issue. Plume visibility models use essentially the same formulations as do single-source Gaussian plume models to describe plume dispersion; thus, some of the problems and uncertainties associated with air quality models in general apply equally to plume visibility models.

However, there are important differences between modeling visibility and modeling ground-level concentrations that should be noted. Visibility modeling is not concerned with ground-level concentrations but with integrals of plume concentrations along the line of sight and generally through the plume center. Because these lines of sight are generally nearly horizontal, it becomes very important to specify accurately the vertical thickness (σ_z) of plumes; the specification of horizontal width (σ_y) is less important. Also, a significant visual impact can result from an elevated plume whose ground-level impact is zero, and from plumes that are tens of kilometers away from the observer, particularly in the clean areas of the western United States.

Atmospheric chemistry is another important model component with significant uncertainties. Modeling the conversion of nitric oxide to nitrogen dioxide is more straightforward than modeling aerosol formation and growth by condensation and coagulation. Therefore, since NO_2 is the principal colorant of power plant plumes, plume discoloration visual effects can be handled much more easily than can visual range reduction from secondary aerosol such as sulfate. Gas phase gas-to-particle conversion can be modeled with relative ease with models like the simple OH• model presented here or with models such as RPM-II (Liu, Stewart, and Roth, 1978). However, the heterogeneous reactions that may be important in humid areas or in plumes with significant concentrations of primary particulate matter are more difficult to model. Finally, aerosol growth (nucleation, coagulation, and condensation) processes must be modeled to account for the time-dependent aerosol size distribution in plumes. Some encouraging reports of work in this case have recently been made (Hobbs et al., 1979; Eltgroth, and Hobbs, 1979; Gelbard and Seinfeld, 1979). Additional theoretical work and corroboration with field measurements is needed in the area of atmospheric chemistry and aerosol dynamics. Some of this work is currently underway at SAI, based on data from the EPA's VISTTA field measurement program in the Southwest.

Additional uncertainties are encountered in modeling the radiative properties of aerosol and gases in the atmosphere. This area has been studied for quite some time. For example, use of the Mie equations to predict particle scatting is now relatively commonplace and has been shown to agree well with empirical data (Latimer et al., 1978). However, one issue that deserves further clarification is the amount of aerosol light absorption in clean areas and in urban atmospheres. Recent studies appear to indicate that aerosol absorption can be quite large (about 40 percent of total extinction) in urban areas and is due to elemental (graphitic) carbon. However, in clean areas the absorption is considerably less (less than 10 percent) and, consequently, relatively unimportant for visibility modeling. If one were to model visibility in urban areas or the effects of urban plumes, however, then absorption would be an important consideration.

After the radiative properties of the pollutants have been determined, the next step is to calculate their effect on atmospheric optics. Atmospheric radiation in the visible spectrum is rather well understood, and sophisticated computer programs are available to compute the multiple scattered radiation. However, multidimensional radiative transfer, such as in a plume, is more difficult to understand and not easily computed. Thus, current visibility models use simplifications to enable the prediction of the effects of pollution. These approximations should be evaluated as to their applicability in specific situations. Currently, the simple approximate technique for computing diffuse radiation in the 1978 version of the SAI visibility model and a more accurate technique are is being evaluated with the EPA's VISTTA data. SAI is currently attempting to develop approximate techniques for handling multiple scattering that will have general applicability and reasonable accuracy.

One of the stated objectives of the EPA's VISTTA program is to obtain sufficient field data so that the accuracy of the SAI plume visibility model can be assessed by comparison with measurements. This model evaluation work is currently underway, and preliminary results are quite encouraging. However, the scope of the VISTTA model evaluation tasks are limited chiefly to visual effects within the first few tens of kilometers of emissions sources. Additional model evaluation efforts are needed for impacts at distances on the order of 100 km from emissions sources. Also, considerable work is needed to refine and evaluate regional models, particularly for applications in complex terrain.

Reliable model predictions of haziness (visual range reduction) caused by regional emissions will probably not be possible within the next 5 years. Therefore, within the next several years regulatory use of visibility models should be limited to plume visibility model applications, where plume discoloration within about 100 km of an emission source is the principal concern.

We conclude with a reiteration of the recommendations made in the first section of this paper, namely that visibility regulations should be designed so that they are both effective and easily enforceable. The possiblity of using existing NSPS and PSD regulations alone to protect visibility should be studied further. Alternatively, visibility regulations could be written in terms of surrogate measures such as plume NO_2 and fine particulate concentrations. If these simpler regulatory strategies are found to be ineffective and visibility standards are written in terms of optical parameters, we strongly recommend that standards be based on a firm set of criteria and on detailed analysis of the aesthetic benefits of good visibility weighed against the costs of restoring and protecting visibiilty.

ACKNOWLEDGMENT

Part of this work was funded by the EPA under Contract 68-02-3337. We gratefully acknowledge Mr. Steven L. Eigsti, the EPA project officer on this contract, for helpful comments and suggestions. The recommendations and conclusions in this paper are the authors' and do not reflect government policy.

REFERENCES

Altshuller, A. P., 1979: Model predictions of the rates of homogeneous oxidation of sulfur dioxide to sulfate in the troposphere. Atmos. Environ. 13, 1653-1661.

Bowers, J. F., 1979: Interim techniques for assessing power plant visibility impacts. Paper presented at the Air Pollution Control Association Specialty Conference on Visibility, 27-28 November 1979, Denver, Colorado.

Calvert, J. G., F. Su, J. W. Bottenheim, O. P. Strausz, 1978: Mechanism of the homogeneous oxidation of sulfur dioxide in the troposphere. Atmos. Environ., 12, 197-226.

Carpenter, S. B., et al., 1971: Principal plume dispersion models: TVA power plants. J. Air Poll. Contr. Assoc., 21, 491-495.

Eltgroth, M. W., and P. V. Hobbs, 1979: Evolution of particles in the plumes of coal-fired power plants - II. A numerical model and comparisons with field measurements. Atmos. Environ., 13, 953-975.

Federal Register, 1979: Protection of visibility; advance notice of proposed rulemaking. 69116-69119.

EPA, 1979: Protecting visibility: An EPA report to congress. Preprint, October 1979, EPA-450/5-79-008.

Gelbard, F., and J. H. Seinfeld, 1979: The general dynamic equation for aerosols, theory and application to aerosol formation and growth. J. Colloid Interface Sci., 68, 363-382.

Hampson, R. F., Jr., and D. Garvin, 1978: Reaction rate and photochemical data for atmospheric chemistry-1977. NBS Special Publication 513, National Bureau of Standards, Washington, D.C.

Hayes, S. R., 1979: A technique for plume visualization in power plant siting. J. Air Poll. Contr. Assoc., 29, 840-843.

Hobbs, P. V., D. A. Hegg, M. W. Eltgroth, and L. F. Radke, 1979: Evolution of particles in the plumes of coal-fired power plants - I. Deductions from field measurements. Atmos. Environ., 13, 935-952.

Isaksen I.S.A., E. Hesstredt, and O. Hov, 1978: A chemical model for urban plumes: Test for ozone and particulate sulfur formation in St. Louis urban plume. Atmos. Environ., 12, 599-604.

Latimer, D. A., 1978: Power plant impacts on visibility in the west: Siting and emission control implications. J. Air Poll. Contr. Assoc., 30, 142-146.

Latimer, D. A., 1979: Power plant impacts on air quality and visibility: siting and emission control implications, EPA-230/11-79-001, Systems Applications, Incorporated, San Rafael, California.

Latimer, D. A., et al., 1978: The development of mathematical models for the prediction and anthropogenic visibility impairment. EPA-450/3-78-110a, b, and c, Systems Applications, Incorporated, San Rafael, California.

Latimer, D. A., and G. S. Samuelsen, 1978: Visual impact of plumes from power plants: A theoretical model. Atmos. Environ., 12, 1455-1465.

____________, 1975: Visual impact of plumes from power plants. UCR-ARTR-75-3, UCI Air Quality Laboratory, School of Engineering, University of California, Irvine, California.

Liu, M. K., D. A. Stewart, and P. M. Roth, 1978: An improved version of the reative plume model (RPM-II). Ninth NATO/CCMS International Technical Meeting on Air Pollution Modeling, 28-31 August 1978, Toronto, Canada.

Miller D. F., 1978: Precursor effects on SO_2 oxidation. Atmos. Environ. 12, 273-280.

Norberg, C. A., R. B. Husar, and D. E. Patterson, 1978: Light scattering characteristics of an aging power plant plume. Third Conference on Atmospheric Radiation, 28-30 June 1978, Davis, California.

Ozkaynak, H., R. G. Isaacs, and B. L. Murphy, 1979: Sensitivity analysis for models of local and regional visibility degradation. Proc. of the Fourth Sumposium on Turbulence, Diffusion, and Air Pollution, 15-18 January 1979, Reno, Nevada, 257-268.

Singh, H. B., F. L. Ludwig, and W. B. Johnson, 1978: Tropospheric ozone: Concentrations and variablities in clean remote areas. Atmos. Environ., 12, 2185-2196.

Sorge, J. M., and V. A. Mohnen, 1979: Visibility modeling using radiative transfer techniques. Proc. of the Fourth Symposium on Turbulence Diffusion and Air Pollution, 15-18 January 1979, Reno, Nevada, 269-272.

Turner, D. B., 1970: Workbook of Atmospheric Dispersion Estimates. Office of Air Programs Publication No. AP-26, U. S. Environmental Protection Agency, U.S. Government Printing Office.

Whitten, G. Z., J. P. Killus, and H. Hogo, 1980: Modeling of simulated photochemical smog with kinetic mechanisms. EPA 600/3-80-002a, Systems Applications, Incorporated, San Rafael, California.

Williams, M. D., M. J. Wecksung, and E. M. Leonard, 1978: Computer simulation of the visual effects of smoke plumes. Society of Photo Optical Instrumentation Engineers Technical Symposium, 30 March 1978, Washington, D.C.

APPENDIX--OTHER SOURCES OF OH·

Two other possible sources of OH· deserve mention, but are not included in the chemical model discussed here. The first of these sources is the photolysis of formaldehyde and related compounds:

$$HCHO + h\nu \rightarrow H\cdot + HCO\cdot \quad , \qquad (A\text{-}1)$$

$$H\cdot + O_2 \rightarrow HO_2^{\cdot} \quad , \qquad (A\text{-}2)$$

$$HCO\cdot + O_2 \rightarrow CO + HO_2^{\cdot} \quad , \qquad (A\text{-}3)$$

$$HO_2^{\cdot} + NO \rightarrow OH\cdot + NO_2 \quad . \qquad (A\text{-}4)$$

Formaldehyde is generated by the catalytic oxidation of hydrocarbons. The catalyzing agent is the hydroxyl radical, which oxidizes hydrocarbons to produce aldehydes but which also oxidizes the aldehydes. The result is an aldehyde equilibrium that is some fraction of the ambient hydrocarbon concentration, generally from 5 to 20 percent depending on the reactivity of the hydrocarbons involved. Methane is an exception: The clean air background methane concentration of 1.5 ppm results in a minimum aldehyde concentration of less than 1 ppb.

In urban atmospheres, where aldehyde concentrations higher than 0.1 ppm have been observed, aldehyde photolysis is the dominant source of hydrogen-containing radicals such as OH·. However, in rural atmospheres hydrocarbon concentrations and reactivity are much lower than in polluted urban air. A typical rural concentration of 0.03 ppmC nonmethane hydrocarbon would be expected to result in only 2 to 4 ppb formaldehyde. This represents a potential hydroxyl source that is 15 to 30 percent of that to be expected from ozone photolysis, assuming a 40 ppb background concentration of ozone. At high plume concentrations (and low ozone) formaldehyde photolysis might be expected to be the principal source of hydroxyl radicals. However, formaldehyde regeneration by hydrocarbon oxidation would be suppressed owing to the very low hydroxyl concentrations in high NO_x environments. Thus, the background concentration of formaldehyde would be depleted before appreciable oxidation to nitrate and sulfate could occur.

For dilute cases there is reason to believe that the 15 to 30 percent contribution from aldehyde photolysis is an overestimate. For aldehyde photolysis to hydroxyl, Reaction (A-4) must react hydroperoxy radicals with nitric oxide (NO). At low concentrations of NO_x, however, nitric oxide is at a very low concentration, and hydroperoxy radicals may build up and react to form hydrogen peroxide:

$$HO_2^{\cdot} + HO_2^{\cdot} \rightarrow H_2O_2 + O_2 \quad . \qquad (A\text{-}5)$$

Because of the low expected yield of hydroxyl from aldehyde photolysis and because of the complexities involved in its treatment, this source of hydroxyl radicals is ignored at this time.

The other potential source of hydroxyl deserving mention at this point is:

$$NO_2 + NO + H_2O \rightarrow 2\ HONO \quad , \qquad (A\text{-}6)$$

$$HONO \rightarrow NO + OH\cdot \quad . \qquad (A\text{-}7)$$

Reaction (A-6) has been shown to proceed very slowly in the gas phase. If Reactions (A-6) and (A-7) are a significant source of OH·, then Reaction (A-6) must have a large heterogeneous component. There exists no clear methodology for treating such a reaction scheme at this time (see Hampson and Garvin, 1978, and Whitten, Killus, and Hogo, 1980).

Appendix I

WET AND DRY REMOVAL, RESUSPENSION, AND ATMOSPHERIC CHEMISTRY OF SO_x AND NO_x

(prepared by D. Randerson)

1. INTRODUCTION

As crude oil and natural gas become increasingly expensive, other less costly fuels will be adopted. Based on recent estimates there are approximately 10^{11} tons of minable coal in the United States. The availability and accessibility of this coal makes it one of the more attractive candidates for coping with our energy demands. However, it is clear that the use of coal (and even oil from shale) poses a number of environmental problems. Two of the most severe are associated with the mining and burning of coal. When coal (or oil) is burned, high concentrations of sulfur and nitrogen oxides as well as smaller amounts of CO_2, CO, and fly ash are released to the atmosphere. The oxides of sulfur and nitrogen are precursors of strong acids and are transformed to compounds that have been associated with visibility degradation, acid rain, health effects, weather and climate effects, and reduction in air quality. Presently, we possess only a limited understanding of the atmospheric processes that lead to these effects. It is the purpose of this section to summarize this understanding as it pertains to deposition, washout, resuspension, and chemical change in relation to sulfur and nitrogen. In the discussion that follows, the principal oxide of sulfur is sulfur dioxide (SO_2), and the important oxides of nitrogen (NO and NO_2) are combined as NO_x.

1.1 Effluent Pathways

When SO_2 and NO_x are released to the atmosphere, they are immediately acted on by various physical and chemical processes. Figure 1 helps portray these processes. If the gases are warmer and lighter than the ambient air, they will rise buoyantly to an equilibrium level. At this level the wind can transport the effluents great distances. As the gases travel with the wind, turbulence mixes the ambient air and the effluents and thereby disperses the gases vertically and horizontally. In the presence of solar radiation, water vapor, and other gases, a myriad of chemical reactions are possible, transforming the SO_2 to sulfate compounds and the NO_x to nitrates. Precipitation scavenging, dry deposition, impaction, and soil and foliar absorption account for the removal of some of the gases and of the transformed material from the atmosphere. These atmospheric mechanisms are responsible for delivering effluent material to the ground and to vegetation;

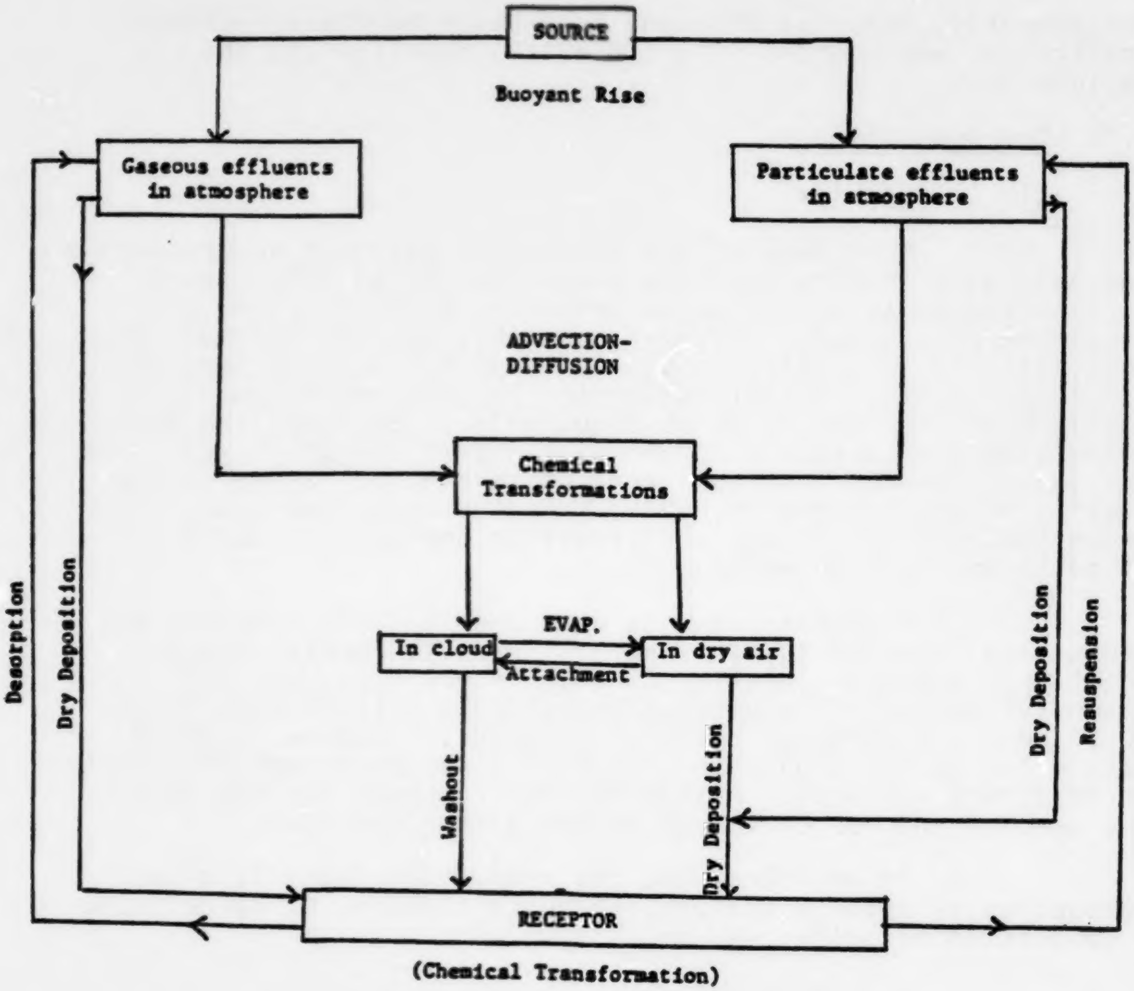

SOURCE
Buoyant Rise
Gaseous effluents in atmosphere
Particulate effluents in atmosphere
ADVECTION-DIFFUSION
Chemical Transformations
EVAP.
In cloud
In dry air
Attachment
Desorption
Dry Deposition
Washout
Dry Deposition
Dry Deposition
Resuspension
RECEPTOR
(Chemical Transformation)

FIGURE 1 Simplified Schematic of Potential Effluent Pathways between Source and Receptor

consequently, they are of prime importance to the assessment, prediction, and control of air quality, both locally and regionally.

1.2 Dry Deposition

1.2.1 Overview

Knowledge of the deposition rates of anthropogenic and natural effluents (gaseous and particulate) onto ground-level surfaces is essential in determining the lifetime of atmospheric compounds, the distances they are transported, and their airborne concentrations. For example, without removal processes, sulfur concentrations in the atmosphere could increase at the rate of about 70 $\mu g/m^3/yr$. By comparing this value with the current U.S. annual standard (40 $\mu g/m^3/yr$), one can comprehend the importance of removal processes to the quality of the air (Hales, 1978). Furthermore, the dry deposition of SO_2 and NO_x contribute to the acidification of soils and surface waters.

Many measurements have been made to document the deposition rates of SO_2 in view of the environmental concern about this species. Similar concern is mounting over the potential health and environmental effects attributable to NO, NO_2, and their chemical reaction products. However, limited data exist on the deposition rates of nitrogenous air pollutants on to ground surfaces. Available data for both the SO_2 and NO_x species and for particulates are summarized here.

In practice, the dry deposition flux, F, of an aerosol is related to the ambient concentration, C, by defining a deposition velocity, v_d, as

$$v_d = F/C$$

In this way dry removal is conveniently parameterized for modeling purposes. Here, v_d has the units of L/T and is usually reported in units of cm/sec while F is expressed as M/L^2T.

1.2.2 Measurements

Methods used for the measurement of the dry deposition of gases appear to fall into three categories (Garland, 1978): 1) the concentration gradient method, 2) tracer methods, and 3) mass balance methods. Another method, the eddy correlation procedure may hold some promise for future application.

The concentration gradient method has been widely used over grasses, soil, water, and wheat. It is based on an expression of the form, $F = k(z)dC/dz$. This method requires simultaneous measurement of the gradient of the gas concentration with height, z, and the appropriate micrometeorological variables to calculate the eddy diffusivity, K. The method can give accurate results only when measurements are made in the constant flux layer over an extensive horizontally homogeneous surface. It cannot be used to estimate uptake by hedges, trees, and other irregularly shaped landscape objects. Normally, measurements of gaseous concentration, temperature, and wind speed are made a few meters off the ground.

Compounds deposited during a brief exposure can be distinguished from natural compounds present in soil or vegetation by tracer methods. This method is well-suited to laboratory experiments and small-scale field experiments, but it has the disadvantage that contamination of the surface prevents a succession of measurements over the same area.

Mass balance experiments have been conducted in the laboratory to measure the rate of removal of SO_2 from the air in an enclosure. The method has also been used in the field to observe the accumulation of sulfur in pots of soil with or without vegetation or in the leaves of trees exposed to elevated SO_2 concentrations. Each field experiment can give rise to difficulties in interpretation when the results are used to estimate the rate of removal of SO_2 from the atmosphere. These difficulties can result because pot plants are more openly exposed than plants in a crop canopy, and they may receive some sulfur from rain. In addition, tree leaves may lose sulfur by translocation to other parts of the plant or by the washing action of rain.

1.2.3 Dry deposition velocity of gases+

A summary of the results of laboratory and field experiments is difficult to construct due to the many factors affecting the estimates. Some of these factors are experimental method, type of surface, reference height, moisture conditions of the soil, and atmospheric conditions (McMahon and Denison, 1979). Nevertheless, some useful information can be gleaned from the data.

+Much of the material in this section as well as in Sections 1.2.4 and 1.3 was extracted from Atmospheric Environment 12:1-3, 1978, pp. 7-23 and from the many papers contained in this volume.

Over dry or wet soil, deposition velocities for SO_2 are of the order of 0.1 to 1.0 cm/sec (Garland, 1978). In fact, studies in Europe demonstrate that SO_2 is removed by dry deposition with an average velocity of 0.8 cm/sec for most surfaces of practical significance. This value probably is correct to within a factor of two when used to provide an estimate of annual deposition over flat, open country dominated by grass. Deposition velocities over heavy vegetation range from 0.1 to 4.0 cm/sec, but a typical value is about 0.6 cm/sec. Larger values seem to occur over forests, but results are still uncertain.

A midday maximum in v_d for SO_2 is frequently observed, reflecting the diurnal maxima in stomatal openings and in turbulent mixing. A seasonal variation is also observed with minimum values less than 0.3 cm/sec when vegetation is dry or senescent. The presence of water on foliage can lead to values of 1 cm/sec or greater, providing the pH of the water is greater than 4.

Deposition velocities for SO_2 have been observed over a variety of surfaces. For example, over dry snow, v_d is approximately 0.1 cm/sec; however, Barrie and Walmsley (1978) have reported geometric means of 0.3 to 0.4 cm/sec in northern Canada. Sparse experimental evidence suggests that snow is a very poor sink for SO_2. Over water v_d ranges from 0.2 to 1.5 cm/sec with a typical value of 0.7 cm/sec. Countryside values of v_d range from 0.2 to 2.0 cm/sec with typical values of 0.8 cm/sec. Observations in London during a prolonged stagnant period suggested a v_d of about 0.7 cm/sec, but caution is advised in using this value. Laboratory measurements of deposition velocities for NO range from 0.1 to 0.2 cm/sec and for NO_2, 0.3 to 0.8 cm/sec (Judeikis and Wren, 1978). These results are still tentative.

1.2.4 Dry deposition velocity of particles

The deposition velocity of particles is strongly dependent on particle size. In general, most sulfate mass lies in the range 0.1 to 1.0 μm diameter for which the deposition velocity is not expected to exceed 0.1 cm/sec. In fact, field measurements suggest that a minimum deposition velocity of the order of 0.01 to 0.1 cm/sec can be expected for particles with diameters of 0.1 to 1.0 μm. Large particles (diameter $>$ 30 μm) fall to the ground under the influence of gravity. Particles smaller than 30 μm are influenced strongly by turbulent motions in the air and deposit at a rate considerably greater than predicted on the basis of gravity alone. At very small sizes ($<$0.01 μm) the deposition process becomes increasingly like that of a gas, with a strong component directly related to the diffusivity.

The deposition velocity of particles is a function of atmospheric stability, roughness length, friction velocity, and particle density. In fact, v_d values for particles are likely to be site specific. Consequently, considerable care must be used in choosing "typical" values of v_d for particles.

In general, the dry deposition of particles can be estimated to within a factor of 10 out to distances of 10 km. Beyond 100 km estimates can be made to within two orders of magnitude.

1.3 Washout (Wet Deposition) for Gases

The relative importance of dry and wet deposition is rather uncertain because of difficulties in obtaining precise measurements of these two processes. However, sulfur budget studies suggest that 50 to 90% (Hales, 1978) of the atmospheric sulfur burden is removed by wet processes (acid rain). The higher of these estimates were based on radioactive fallout measurements and can be expected to overestimate the wet component for near-surface sources. Moreover, some of the more specific studies have indicated dry deposition to be of relatively greater importance. Regardless of these features, there seems to be little doubt that precipitation scavenging constitutes an important, and in some cases, the dominant mechanism for sulfur removal.

The washout of a soluble gas involves gas and liquid phase mass transfer rates and the solution equilibria of the gas. For example, with SO_2 these solution equilibria involving sulfurous acid are pH dependent, and low pH suppresses the solubility of SO_2 and the ionization of the acid. The washout of SO_2 by rain falling through a uniform concentration of SO_2 is a function of: 1) the size spectrum of the droplets and hence the rainfall rate, 2) the initial pH of the rain, and 3) the absolute magnitude of the SO_2 concentration. In general small rain drops (<1 mm in diameter) are saturated with SO_2 while larger drops do not achieve saturation. Typical concentrations of sulfate in precipitation range from 10 to 100 μmole/liter with much higher values occurring in specific samples occasionally (Hales, 1978).

At rainfall rates of ≤ 2 mm/hr the liquid phase resistance to mass transport is insignificant (Marsh, 1978) so that solubility equilibrium exists. Although equilibrium scavenging conditions generally do not occur in the vicinity of concentrated SO_2 sources, they do prevail at moderate distances downwind and can be accepted for the purposes of regional SO_2 washout calculations.

According to Marsh (1978), the dominant ions contributing to the acidity of precipitation are usually SO_4^{2-}, NO_3^-, and NH_4^+ after allowing for sea salts. In remote areas with low concentrations of SO_2, theoretical estimates suggest that the washout of the gas is unlikely to produce marked changes in the pH of rain although the presence of NH_3 can enhance the final sulfate concentration. In general, sulfate removed by rain has a time constant of about $10^{-4} sec^{-1}$. Theoretical calculations indicate that the half-life of SO_2 removal from the atmosphere at a 1 mm/hr rainfall rate (with a well-mixed boundary layer 1 km deep) ranges from several hours to days, depending on the background pH and initial atmospheric SO_2 concentrations.

In the eastern United States and in western Europe about 2/3 of acidity in precipitation is due to SO_2 and 1/3 to NO_x, although the proportion of the acidity due to NO_x has been increasing in recent years. By contrast, in the western United States, 2/3 of the acidity is due to NO_x because the automobile contributes more to the pollution burden than do power plants and heavy industry.

1.3.1 Washout for particles

Based on limited data, an approximate expression for the rain scavenging rate for particles is given by (see Slinn, 1977)

$$\psi_r(a) = P(\vec{r},t)\,E(a,R_m)/2R_m \qquad (1)$$

where ψ_r is the rain removal rate $[T^{-1}]$, a is the radius of the aerosol particle, $P(\vec{r},t)$ is the rainfall rate [L/T] at vector location $\vec{r}$ and time t, E is the collision efficiency between drops and particles and R_m is the volume-mean drop radius as $R_m = 0.35\ mm(p/1\ mm/hr)^{1/4}$.

Theoretical considerations and laboratory observations indicate that the collision efficiency is a function of mean drop size, and it ranges from approximately 0.1 for $a = 10^{-3}\ \mu m$ to 5×10^{-4} for $a = 0.1\ \mu m$ to near 1.0 for $a = 10\ \mu m$. Experimental data suggest that ψ_r is proportional to $p^{3/4}$. However, field data show that particle growth in clouds can be substantial. Predictions of precipitation scavenging can be made to within a factor of 2 to 4 if not too much demand is placed on the spatial and temporal resolution required, and if there is information available on the particle size distribution (Slinn, 1981).

Crude estimates of the scavenging of particulates by ice crystals can be obtained from an expression similar to Eq.1. This expression yields a snow removal rate, ψ_s , as a function of the particle/ice crystal collection efficiency, E, and the snowfall rate (rain equivalent), p(mm/hr). For example, for a snowfall-rain equivalent rate of 1mm/hr, $\psi_s = 10^{-4}$/sec for rimed crystals.

The possible range of wet deposition for particles is from 0 (no precipitation) to essentially complete removal where all particles could be deposited within approximately 1000 sec. In this range of possibilities, the probability of a specific deposition pattern depends on the rate of atmospheric diffusion of the effluent, the probability of precipitation at the time the effluent is over a given area, and on the probability that the wind is advecting the material toward the area of concern.

1.4 Washout Coefficients and Ratios

The washout of a gas by rain is usually described in terms of the washout coefficient λ, defined as the fraction of the gas removed in unit time by rain below the cloud base. Washout coefficients for various gases and particulates depend on many factors such as raindrop size, pH, temperature, effluent concentration, duration and intensity of rainfall, and aerosol characteristics. Even though many factors contribute to the value of λ, there seems to be some consistency among λ values calculated from field observations. For example, McMahon et al. (1976) and Maul (1978) suggest that for SO_2

$$\lambda = \alpha R \times 10^{-5} \text{ sec}^{-1}$$

where α is a constant of about 3 and R is the rainfall rate in mm/hr. In general, λ for SO_2 appears to range in value from 10^{-7} sec^{-1} for light snow ($R < 0.1$ mm/hr) to 10^{-4} during heavy rainshowers or thunderstorms ($R = 10$ mm/hr). Estimates of λ for NO_x are approximately 25% of those of SO_2 ($\lambda[NO_x] = 0.25\,\lambda[SO_2]$). For particles $\lambda \simeq 3 \times 10^{-5}$ for $R = 1$ mm/hr and a particle diameter of 1 um. For smaller particles λ appears to get smaller ($\lambda \simeq 10^{-7}$ for 0.1 μm particles).

For wet deposition processes there is considerable support for the "scavenging or washout ratio" approach. This method has had a large amount of exposure in recent years, mainly as a consequence of studies of the wet deposition of trace metals. In essence, the washout ratio is the ratio of the concentration of material detected in precipitation to that measured in air at the same time. In general, the conventional washout ratio for SO_2 on a volume basis is approximately 5×10^4 (e.g., see Henmi and Reiter, 1978). Scavenging ratios appear to be a function of particle size and rainfall intensity. Generally, these ratios decrease with increasing precipitation rates and also increase with particle size. In addition, the ratios appear to increase with distance from the source.

1.4.1 Wet deposition velocity parameterization

The concept of a wet deposition velocity (v_w) can be used for both particles and gases. For particles, this parameterization is always appropriate, regardless of the time and space scales of interest. For particles, $v_w = w_r R$ where w_r is the washout ratio and R is the precipitation rate. The wet deposition velocity parameterization for gases, however, should be used cautiously unless the plume is well mixed in the atmospheric boundary layer. When it is, then for gases that form simple solutions in water, $v_w = \theta R$, where θ is an appropriate solubility coefficient. For example, θ for SO_2 at 18°C is 42, for CO_2 at 18°C it is 0.95, for H_2S at 10°C it is 3.4, and for NH_3 at 25°C it is 450. Solubility coefficients typically decrease with increasing temperature and usually decrease with the presence of other constituents. For reactive gases, $v_w = \theta_* R$, where θ_* is the appropriate enhanced solubility coefficient. For SO_2, θ_* is of the order of 10^4.

1.5 Gas-to-Particle Conversion Rates

A topic of obvious interest in air quality analysis is the rate of conversion of trace gases like SO_2 and NO_x to the solid or liquid state. The products of this gas-to-particle conversion process are thought to play an important role in visibility degradation and in the production of acid rain. Sulfur dioxide has been intensely studied in this regard because it is a likely precursor for ubiquitous sulfate particles in the atmosphere. More recently there has been increasing interest in atmospheric nitrates which could be produced by the oxidation of NO_x.

Because of the interest in sulfate particles, the SO_2-to-SO_4^{2-} conversion process has been studied in power plant plumes and in urban plumes. Although there is still disagreement as to the nature of the predominant conversion mechanism, there is general agreement as to the magnitude of the conversion rate. In general, in power plant plumes the rate of conversion of SO_2 to sulfate is from 0.2 to 2.0% hr^{-1} (Hobbs *et al.*, 1979) or from 0.5 to 5.0% (Forrest and Newman, 1977). Values of 1% or more tend to be associated with increased travel time, high insolation values, high temperatures, and in some cases high relative humidity. For example, during midday summertime conditions, SO_2 conversion rates in power plant plumes can reach 1 to 2% hr^{-1}. This rate can be explained in terms of the OH induced oxidation of

SO_2 (Davis et al., 1979, and Chang, 1979). Under very favorable conditions the rate might increase to 4 to 6% hr^{-1} (Hegg and Hobbs, 1980). In fact, field data indicate that as a power plant plume mixes with ambient air the rate of conversion to sulfates increases. Thus, tall stacks reduce ground-level concentrations of SO_2 but contribute to increasing sulfate aerosol formation by reducing surface losses of SO_2 and by increasing the atmospheric residence time, thereby increasing the SO_2-to-SO_4^{2-} conversion. Here the implications for acid rain are obvious. The conversion $SO_2 + OH$ to a sulfate takes place at a rate of about 1.6×10^3 ppm^{-1} min^{-1} (Williams et al., 1980).

In urban plumes the conversion rate tends to be faster--1 to 10% hr^{-1} (Forrest et al., 1979). However, these values are the subject of much debate. Fast conversion rates in urban plumes seem to be related to high insolation values.

The conversion rate of NO_x to nitrates is less well documented. Preliminary results from field projects (e.g., Hegg and Hobbs, 1979) point to rates of 0.1 to 0.5% hr^{-1} with rates as high as 4% hr^{-1} over urban areas during daylight. The data also suggest an increase in conversion as a function of travel time and of time of day. For example, Davis et al. (1979) suggest that during midday summertime conditions, conversion of NO_2 to HNO_3 may occur very early in the chemical life of a power plant plume (3-5 hr). The consequences of this condition could result in HNO_3 being the major acid species present near the source. The conversion $NO_2 + OH \rightarrow HNO_3$ takes place at about 9×10^3 ppm^{-1} min^{-1}.

1.6 Residence Times

An important end product of the analysis of the atmospheric cleansing process is to predict and interpret the atmospheric residence (e-fold or turnover) times for various effluents. Since the contribution to the removal of a pollutant by washout is on a time scale comparable to the atmospheric water cycle (1 wk), any analysis of residence time must involve global scale-space averages. With this in mind the residence time, t, is given by $t^{-1} = (t_d + t_w + t_c + t_p)^{-1}$ where the t's are the removal times for dry, wet, chemical, and physical processes, respectively. For example, if the characteristic length scale of the mixed layer is 1 km and v_d is 0.8 cm/sec, then t_d is approximately 40 hr. Consequently, if the average wind speed were 25 km/hr, then a mean transport distance of 1000 km is possible before deposition occurs. In general, Henmi and Reiter (1978) found that during summertime over the eastern United States the residence time of SO_2 due to dry deposition appears to range from about 50 hr on the east coast to near 70 hr along 105°W. In winter the corresponding range is from near 30 hr to near 40 hr along 105°W. The residence time of SO_2 due to wet removal in the bottom km of the atmosphere is of the order of 50 hr. However, by restricting the calculations to the mixed layer (depth, 1 km) a gross underestimate of residence time is made. Residence times for tropospheric aerosols are of the order of 7 to 10 days. Consequently, transport distances of the order of 4000 to 6000 km may occur.

Some preliminary estimates of the residence time of NO_x over urban areas are available. Results point to residence times of two days or less (Noxon, 1975).

1.7 Resuspension[+]

Resuspension describes the process by which material deposited on the ground is subsequently reentrained or resuspended into the atmosphere. Resuspension occurs both as a result of the action of wind on the ground and of any mechanical disturbance of the soil. Mechanical stresses such as digging or walking on soil can lift particulate material above the ground and into the main airstream where the particles are carried away by the wind and mixed upward by mixing processes. The plowing of fields and other agricultural practices also act as an important source for particulate matter.

In air-surface mass transfer processes, the most important particles to consider for both deposition and resuspension are those that can be inhaled and potentially cause damage to the lungs. The respirability of a particle depends on its diameter. Particles less than 3.5 μm in diameter are most often considered respirable, but some data indicate that particles as large as 10 μm can be inhaled.

From the aspect of visibility, the important impacts are caused by the scattering of light from particles between 0.1 and 1 μm in diameter (and by the absorption of light by NO_2). Consequently, from an air quality viewpoint, the important part of the particle size spectrum includes particles with diameters of 0.1 to 10 μm.

Resuspension rates can be used to estimate the flux of particulates into the main airstream. A resuspension rate is defined as the vertical flux of particulate matter per unit area divided by the particulate concentration. Resuspension rates appear to vary from 10^{-12} to 10^{-4} fraction resuspended/sec. Most experimental values lie in the range

---

[+] Material in this section was drawn from the review by Sehmel (1981).

10^{-9} to 10^{-7} $\sec^{-1}$. Resuspension rates from vehicular traffic range from 10^{-5} to 10^{-2} $\sec^{-1}$. Clearly, with the uncertainties in this rate, more experimental data is needed before much confidence can be placed on estimates of air quality impacts from resuspension.

1.8 Comments

Many uncertainties exist about washout processes for particles and gases. For particles there are major uncertainties about the retention efficiency, about particle modifications in the atmosphere, and about the collision efficiency for particles of 0.1 to 1.0 μm in diameter. There are so little data on snow scavenging of particles that even the uncertainties are unknown. For gases there are major uncertainties about adequate parameterizations for reaction rates, and there is also the problem of determining the amount of gas that attaches to aerosol particles.

To predict the short-term local deposition and washout of gases and particles is a difficult task. The primary difficulty arises from problems related to forecasting the details of the meteorological conditions, especially the precipitation rate, aerial extent, and type. Remember, the range of wet removal is from zero to nearly complete removal. Another problem for short-term releases is knowledge of the chemical properties of the suspended material and its vertical distribution in the atmosphere.

Long-term, average removal, out to distances of about 100 km is simpler to describe or to predict than that for a short-term, accidental release because average meteorological data can be used. For example, for seasonal average predictions of, say, the transport-diffusion-wet deposition of particles, it is consistent to use a Gaussian plume model to describe the transport diffusion process and to multiply by a simple decay term ($\exp[-x\bar{\Lambda}/\bar{u}]$) to describe the decrease in particle concentration caused by washout. Also, removal rates can be parameterized in terms of removal rates or wet deposition velocities. Indeed, regardless of initial particle size, it appears to be an accurate estimate to specify that the efficiency with which particles are removed by a storm, ϵ, is the same as the efficiency with which the storm removes its condensed cloud water. For gases that form simple solutions in water and for problems at large distances, it is almost certainly acceptable to set the gas concentration in surface level rain, k_0, equal to the value in equilibrium with the surface level air contamination, x_0; namely, $k_0 = \beta x_0$ where β is the solubility coefficient.

1.9 Assessing Air Quality

Our ability to assess air quality over large regions depends on our capabilities to solve the transport-diffusion-deposition-resuspension-transformation problem. Errors introduced in estimating each of these modules contributes to the total error in the simulation of atmospheric concentrations and thus to any objective evaluation of air quality. As noted in the sections on the transport-diffusion modules, accuracy to within a factor of 2 (e.g., see American Meteorological Society, 1978) has been estimated frequently in connection with routine applications of dispersion modeling. When research-grade instruments are used, accuracies of 10 to 40% can be expected over short distances ($\leq$10km). For the deposition-resuspension modules the accuracies are slightly less than those of the transport diffusion module. If not too much demand is placed on the spatial and temporal resolution, then predictions of reality can be made to within a factor of 2 to 4. However, for spatial and temporal resolutions appropriate for an accidental release, even order-of-magnitude estimates are elusive. These errors represent the inherent accuracy of the mathematical models used to estimate air quality. Consequently, decision on air quality impacts can only be improved if these uncertainties are correctly taken into account.

At present the National Ambient Air Quality Standards (NAAQS) do not effectively address the long-range transport of compounds such as sulfates and nitrates. These two species are important components of the criteria mixture entitled total suspended particulates (TSP). However, since they are emitted primarily as gases and because of their fine particulate size, they are not controlled to a meaningful extent under TSP regulations. In addition, the deposition (wet and dry) of sulfates and nitrates is not viewed as a violation of existing air quality standards.

Power plant and urban plumes have been detected at distances of 500 to 1000 km from their sources. Consequently, SO$_2$ generated in power plant plumes and O$_3$ generated from hydrocarbons and NO$_x$ may cause significant increases in pollution far from the source. Air pollution due to secondary pollutants such as sulfates and ozone is not controlled by the government entity where the impact actually occurs. Therefore, current concepts of air quality control regions needs to be reviewed and revised to take into account the long-range transport of secondary pollutants.

1.10 Data Needs and Data Bases

In many cases it is presently more reliable to estimate precipitation chemistry at a specific receptor site on the basis of the deposition climatology for that area than it is to attempt some type of model calculation. However, modeling should never be expected to substitute for monitoring. Furthermore, field measurements in the form of network monitoring are essential for model development. Consequently, a strong interactive relationship must exist between modeling and monitoring activities.

1.10.1 Dry deposition

The following recommendations are made:

1. A reliable, well tested method of measuring dry deposition needs to be developed. Otherwise, develop improved sensors for dry-deposition flux measurement with special emphasis on gradient and covariance techniques.

2. With few exceptions, measurements of dry deposition have been intermittent. More continuous measurements are needed.

3. A regional-scale deposition climatology needs to be developed for use in assesssing trends, in documenting regional air quality variability, and in testing and verifying simulation models. As an alternative, some site-specific studies should be considered.

4. A fundamental requirement is knowledge of the minimum removal rate for, say, particles in the 0.1 to 1.0 μm range. This value would serve as a baseline value for regulatory purposes.

5. Dry deposition is strongly dependent on atmospheric stability and wind speed, especially over wet surfaces. Meteorological measurements of these parameters on a routine network basis should be made near important receptors. In addition, routine network monitoring of SO_2 and NO_2 surface-level ambient concentrations should be taken in conjunction with the meteorological data.

6. Present estimates of the deposition velocity of sulfate varies more than do estimates of SO_2. Many measurements of the deposition of sulfate particles in the 0.1 to 1.0 μm range give velocities of the order of 0.1 cm/sec, about an order of magnitude less than for SO_2. However, some experiments give higher results.

Also, sulfate may be present in larger particles near emitting sources, in sea spray droplets, and in resuspended soil. Estimates of deposition in these circumstances are needed.

7. At high relative humidity, sulfate particles grow by deliquescence and may become large enough to be impacted on surfaces, especially trees. Information is needed on the contribution of fog drip to deposition of sulfate.

8. More dry deposition measurements are needed for forests, cold and hot deserts, and ice and snow surfaces.

9. Compile and maintain a comprehensive North American land use data base for regional-scale dry-deposition calculations. It should be noted here that a land-use inventory conducted by EPA is nearing completion.

10. Continue and improve measurements of dry deposition to natural surfaces under typical atmospheric conditions. Link these studies directly with biological transport studies, when applicable.

1.10.2 Wet deposition

The following recommendations are made:

1. Maintain and upgrade currently available emission-inventory data banks--NO_x and SO_x and other criterion pollutants from both point and nonpoint sources.

2. Encourage rapid completion of the EPA central precipitation chemistry data bank. Make it user accessible.

3. Develop comprehensive case-study data bases for model applications and verification in a unified format including:
    a. complete listing of standard meteorological data,
    b. three-dimensional winds,
    c. precipitation chemistry,
    d. air quality data,
    e. deposition data,
    f. supplemental radar and satellite data.

4. Examine the relative importance of anthropogenic and natural sources of important pollutant species. These should include criterion pollutants as well as ionic species such a NO_3^-, SO_4^{2-}, Ca^{++}, and Cl^-.

5. Define usable relationships for oxidation rates of SO_2 in atmospheric water as a function of pertinent variables (SO_2, O_3, H^+, M^+, organics, etc.).

6. Define and quantify pertinent mechanisms leading to the occurrence of sulfate and nitrate in precipitation.

7. Define the mechanisms affecting the presence of NH_3 in precipitation.

8. Study precipitation scavenging and deposition under conditions of orographic precipitation and in other situations of strong topographic influence, such as the trapping of pollutants in a valley.

9. Perform comprehensive material-balance investigations on representative storm types (convective, cyclonic, orographic) to evaluate scavenging mechanisms and behavior.

10. Perform pilot studies to optimize network density and siting.

11. There is a need for greater time and space resolution in precipitation chemistry sampling, i.e., intrastorm sampling.

12. Establish long-term monitoring networks to supplement the WMO background network with stations on a grid of approximately 200 km, with samples collected for no longer than a month. Consider daily sampling programs or careful intercalibration analyses with shorter-term intensive regional-scale programs.

1.10.3 Data bases

A summary of research activities directly and indirectly applicable to chemical deposition assessment are tabulated in Table A.

1.11 Sulfate Chemistry

When SO_2 is released to the atmosphere, oxidation begins. The mechanisms by which SO_2 is oxidized to sulfates are important because they: 1) determine the rate of formation of sulfate, 2) influence the concentration of SO_2 on the reaction rate, and, to some extent, 3) determine the final form of the sulfate. Various mechanisms have been proposed to account for this oxidation; however, the oxidizing species and pertinent rates of reactions have not been firmly established. Presently, there is not even general agreement as to whether

TABLE A. Some Research Activities Related to Wet and Dry Deposition, Resuspension and Atmospheric Chemistry

Agency or Source	Comments
EPRI	SURE: Sulfate regional experiment
OECD (Europe)	LRTAP: Long-range transport of air pollution
DOC, NOAA	Wet and dry removal modeling, wet removal measurement
DOE	Chemical composition of precipitation, environmental fate of emissions, acid rain, particle resuspension and translocation, air pollution and dry deposition, scavenging of pollutants, nuclear and non-nuclear materials, lab and field programs
DOD	Scavenging modeling and research, collateral damage and chemical warfare
DOT, FAA	NO_x scavenging research
EPA	NEROS: Northeast regional oxidant study, NECRAMP: Northeast corridor regional modeling project, PEPE: Prolonged elevated pollution episode, STATE: Sulfur transformation and transport experiment, MAP3S (transferred from DOE to EPA), LARPP: Los Angeles reactive pollution program (completed)
NASA	Rocket effluent scavenging research
NRC	General assessment modeling
NSF, NCAR	Wet chemistry research and scavenging modeling
TVA	Scavenging and dry deposition measurement and modeling
USDI, USGS	Wet removal analysis

homogeneous or heterogeneous oxidation mechanisms are more effective. A major difficulty in solving this important problem is the paucity of data which are useful in modeling efforts.

Experimental results indicate that SO_2 is oxidized homogeneously to a variety of sulfoxy radicals and compounds by reaction with either of the tropospheric odd hydrogen radicals (OH, HO_2). The two most important reactions (e.g., see Davis et al., 1974, and Kuhlman et al., 1978) for the gas-phase conversion of SO_2 to sulfate are thought to be,

$$OH + SO_2\ (+M) \xrightarrow{k_1} OHSO_2\ (+M) \qquad (R1)$$

$$HO_2 + SO_2 \xrightarrow{k_2} SO_3 + OH \qquad (R2)$$

and the sulfur trioxide (SO_3) is rapidly converted to sulfuric acid by

$$SO_3 + H_2O \xrightarrow{k_3} H_2SO_4 \qquad (R3)$$

In R1 and R2 a rate constant of $k_1 = 10^3$ ppm^{-1} min^{-1} and $k_2 = 1.4$ ppm^{-1} min^{-1} have been measured and $k_3 = 1.3*10^3$ ppm^{-1} min^{-1}. In addition, reactions with free radicals occur,

$$SO_2 + RO_2 \xrightarrow{k_4} SO_3 + RO \qquad (R4)$$

where $k_4 = 0.3$ ppm^{-1} min^{-1}.

In summary, it appears that SO_2 can be oxidized to SO_3 and converted to H_2SO_4 aerosol, or it may form sulfite ions that are then oxidized to sulfate. Subsequent to the oxidation, H_2SO_4 or SO_4^{2-} may interact with other materials to form other sulfate compounds. The most important sulfate formation mechanisms identified to date are summarized in Table B.

1.11.1 Implications to air quality

A key concern for air quality control strategy is to determine whether SO_2 conversion is first order or zero order in SO_2. If the conversion is first order, then an 80% decrease in SO_2 emissions could lead to an 80% decrease in sulfate. By contrast, if the conversion is zero order in SO_2 and if 10% of the emitted SO_2 is converted to sulfate, then a 98% reduction in SO_2 emissions would be required to produce an 80% reduction in sulfate. In the latter case, it is very important to determine if there are other pollutants that influence SO_2 conversion that could be more easily controlled. For example, if a heavy metal catalysis is important (Table B, #5), TABLE B. Mechanism by which SO_2 is converted to sulfates (after Wilson, 1978)

Mechanism	Overall Reaction	Dependent Factors
	Homogeneous Reactions	
1. Direct Photooxidation	$SO_2 \xrightarrow[\text{oxygen}]{\text{light, water}} H_2SO_4$	SO_2 concentration, intensity of sunlight
2. Indirect Photooxidation	$SO_2 \xrightarrow[\text{organic oxidants, } OH,\ HO_2,\ RO_2 \text{ radicals}]{\text{water, } NO_x\text{, smog}} H_2SO_4$	SO_2 concentration, organic oxidant concentration, OH, NO_x, HO_2, RO_2
	Heterogeneous Reactions	
3. Air oxidation in liquid droplets	$SO_2 \xrightarrow{\text{water}} H_2SO_3$ $NH_3+H_2SO_4 \xrightarrow{O_2} NH_4^+ + SO_4^{2-}$	Ammonia concentration
4. Catalyzed oxidation in liquid droplets	$SO_2 \xrightarrow[\text{heavy metal ions}]{O_2\text{, water}} SO_4^{2-}$	Concentration of heavy metal (Fe, V, Mn) ions
5. Catalyzed oxidation on dry surfaces	$SO_2 \xrightarrow[\text{carbon particle}]{O_2\text{, water}} H_2SO_4$	Carbon particle concentration (surface area)

better fly ash controls to limit primary aerosol emissions could be used to reduce sulfate.

1.12 Nitrate Chemistry

Particulate nitrate is one of the important reaction products of nitrogen oxides in the atmosphere. Studies of NO_x reactions suggest that particulate nitrate initially accounts for only a small fraction of the NO_x that is transformed to products under photochemically reactive conditions. However, under less reactive conditions, or during transport of an urban plume, or in nonurban air, particulate nitrate can make up a significant fraction of the oxidized nitrogen (excluding N_2O) in the lower atmosphere.

Gaseous nitrates identified in the atmosphere include nitric acid, peroxyacetyl nitrate (PAN), and peroxypropionyl nitrate (PPN). But even though we know these products exist, our understanding of the distribution and formation mechanisms is not adequate to cope with a potential need to control these products. Actually, very little is known about the processes that convert nitrogen oxides to gaseous and particulate nitrates.

Generation of O_3 in polluted air is determined by the rate at which nitric oxide is converted to NO_2 without the consumption of odd oxygen. This conversion takes place through reactions with HO_2 and organic peroxy radicals; namely,

$$NO + HO_2 \longrightarrow NO_2 + OH \qquad \text{(R5)}$$

$$NO + RO_2 \longrightarrow NO_2 + RO \qquad \text{(R6)}$$

These reactions suggest that the NO_x concentration is important in determining OH and HO_2 radicals. In fact, R5 appears to be the principal source of hydroxyl in urban atmospheres, and it also contributes hydroxyl to R1. The result of this dependence is that the concentrations of OH are directly proportional to those of NO_2 during the midday hours. Consequently, for many photochemical products, the interactions of the hydroxyl radical turn out to be quite important. Such behavior is obvious in the case of nitric acid, which is formed by

$$NO_2 + OH + M \longrightarrow HNO_3 + M \qquad \text{(R7)}$$

in addition,

$$NO + OH \longrightarrow HNO_2 \qquad \text{(R8)}$$

$$HNO_2 + h\nu \longrightarrow NO + OH \qquad \text{(R9)}$$

Reactions 5 and 6 lead to the formation of O_3 through the reactions,

$$NO_2 + h\nu \longrightarrow NO + O \qquad (R10)$$

$$O + O_2 + M \longrightarrow O_3 + M \qquad (R11)$$

and other reactions yield NO_2,

$$NO + O_3 \longrightarrow NO_2 + O_2 \qquad (R12)$$

$$2NO + O_2 \longrightarrow 2NO_2 \rightleftarrows N_2O_4 \qquad (R13)$$

so that the chemical transformation system for NO_x compounds is very complex.

1.12.1 Implications to air quality

Gross estimates are that 60 to 70% of the acid precipitation in the United States is due to sulfuric acids while nitric acids are responsible for the balance of the acidity. However, these proportions are not fixed and speculation suggests that because of the continuing increase in NO_x emission rates, nitric acids will constitute an increasingly larger part of the problem. In addition, products of NO_x and SO_2 oxidation can lead to degradation in visibility. NO_x emissions from power plants and from vehicles can result in discoloration of the sky. Furthermore, the concept of the scale of these problems must be addressed because nitrates and sulfates can be transported great distances and impact areas 100's to 1000's of km from the source. However, our limited knowledge on NO_x and SO_2 transformation processes makes regulation at these distances difficult to formulate and justify. In fact, the compounds most directly responsible for acid precipitation and visibility degradation, sulfates and nitrates, are not directly regulated. As a result, the deposition of sulfates and nitrates cannot be viewed as a violation of existing air quality standards.

1.13 Recapitulation

Over most of the nation's air quality control regions the primary ambient standards for SO_2 and NO_2 have been attained. The problem is that most geographical regions can comply with the ambient standards and still allow substantial emissions of these pollutants. In fact, many states are able to achieve present NO_2 standards by relying solely on federal motor vehicle emission controls, with no requirements in their state implementation plans for NO_2 controls on existing stationary sources. Some suggest that SO_2 control requirements, while more stringent than those for NO_2 are nonetheless far below our technological capability. Furthermore, the standards are especially ineffective in addressing long-range transport because concern for compliance focuses on ground-level concentrations, not on higher altitudes more relevant to the transport process. Consequently, the major pollution control programs for existing sources were not designed to reduce the total atmospheric loading of effluents. Yet it seems that a reduction in the aggregate loading of SO_2 and NO_2 is necessary to reduce acid precipitation and visibility degradation.

As a final point, the international implications of long-range transport and transformation should be considered. Given the great distances that effluents can travel, and the role of both the United States and Canada in contributing to the deposition of acids and to the transport of aerosols across their common border, it is clear that acid rain and visibility degradation are international as well as domestic problems. There is presently no agreement between Canada and the United States concerning transfrontier effluent transport. Hence, international environmental principles offer the central guidance concerning both nations' responsibilities, if any, to ameliorate the problem. Precedents for this type of cooperation have been established through the 1962 Treaty Banning Nuclear Weapons Tests, and perhaps more significantly the formulation of principles and doctrine with regard to the Trail, British Columbia-State of Washington pollution dispute of the 1920's and 1930's.

REFERENCES

American Meteorological Society, 1978, Accuracy of Diffusion Models. Bulletin of the Am. Meteorol. Soc. 58:1025-1026.

Atmospheric Environment, 1978, Sulfur in the Atmosphere, 12, 796 pp.

Barrie, L.A. and J. L. Walmsley, 1978, A Study of Sulphur Dioxide Deposition Velocities to Snow in Northern Canada, Atmos. Environ. 12:2321-2332

Chang, T. Y., 1979, Estimate of the Conversion Rate of SO_2 to O_4 from the DaVinci Flight Data. Atmos. Environ. 13:1663-16664.

Davis, D. D., W. Heaps, D. Philen, and T. McGee, 1979, Boundary Layer Measurements of the OH Radical in the Vicinity of an Isolated Power Plant Plume: SO_2 and NO_2 Chemical Conversion Times. Atmos. Environ. 13:1197-1203.

Davis, D. C., G. Smith, and G. Klauber, 1974, Trace Gas Analysis of Power Plant Plumes via Aircraft Measurement: O_3, NO_x, and SO_2 Chemistry. Science 186:733-735

Forrest, J., S. E. Schwartz, and L. Newman, 1979, Conversion of Sulfur Dioxide to Sulfate During the DaVinci Flights. Atmos. Environ. 13:157-167.

Forrest, J., and L. Newman, 1977, Further Studies on the Oxidation of Sulfur Dioxide in Coal-Fired Power Plant Plumes. Atmos. Environ. 11:465-474.

Garland, J. A., 1978, Dry and Wet Removal of Sulphur from the Atmosphere. Atmos. Environ. 12:349-362.

Hegg, D. A. and P. V. Hobbs, 1980, Measurements of Gas-to-Particle Conversion in the Plumes of Five Coal-Fired Electric Power Plants. Atmos. Environ. 14:99-116.

Hegg, D. A., and P. V. Hobbs, 1979, Some Observations of Particulate Nitrate Concentrations in Coal-Fired Power Plant Plumes. Atmos. Environ. 13:1715-1716.

Henmi, T., and E. R. Reiter, 1978, Regional Residence Time of Sulfur Dioxide over the Eastern United States. Atmos. Environ. 12:1489-1495.

Hobbs, P. V., D. A. Hegg, M. W.Eltgroth, and L. F. Radke, 1979, Evolution of Particles in the Plumes of Coal-Fired Power Plants--I. Deductions from Field Measurements. Atmos. Environ. 12:935-951.

Judeikis, H. S., and A. G. Wren, 1978, Laboratory Measurements of NO and NO_2 Depositions onto Soil and Cement Surfaces. Atmos. Environ. 12:2315-2319.

Kuhlman, M. R., D. L. Fox, and H. E. Jefferies, 1978, The Effect of CO on Sulfate Aerosol Formation. Atmos. Environ. 12:2415-2423.

Marsh, A. R. W., 1978, Sulfur and Nitrogen Contributions to the Acidity of Rain. Atmos. Environ. 12:401-406.

Maul, P. R., 1978, Preliminary Estimates of the Washout Coefficient for Sulphur Dioxide Using Data from an East Midlands Ground Level Monitoring Network. Atmos. Environ. 12:2515-1517.

McMahon, T. A., and P. J. Denison, 1979, Emperical Atmospheric Deposition Parameters--A Survey. Atmos. Environ. 13:571-585.

McMahon, T. A., P. J. Denison, and R. Fleming, 1976, A Long-Distance Air Pollution Transportation Model Incorporating Washout and Dry Deposition Components. Atmos. Environ. 10:751-761.

Noxon, J. F., 1975, Nitrogen Dioxide in the Stratosphere and Troposphere Measured by Ground-Based Absorption Spectroscopy. Science 139:547-549.

Sehmel, G. A., 1981, Deposition and Reinsertion Processes. Atmospheric Science and Power Production (D. Randerson, Ed.), Chapt. 12 (to be published by DOE, TTD--Oak Ridge, TN).

Slinn, W. G. N., 1981, Precipitation Scavenging. Atmospheric Science and Power Production (D. Randerson, Ed.), Chapt. 11 (to be published by DOE, TTD--Oak Ridge, TN).

Slinn, W. G. N., 1977, Some Approximations for the Wet and Dry Removal of Particles and Gases from the Atmosphere. J. Water, Air, and Soil Poll. 7:513-543.

Williams, M. D., E. Treiman, and M. Wecksung, 1980, Plume Blight Visibility Modeling with a Simulated Photographic Technique. J. Air Poll. Control Assoc. 30:131-134.

Wilson, W. E., 1978, Sulfates in the Atmosphere: A Progress Report on Project MISTT. Atmos. Environ. 12:537-547.

Appendix J

TRANSPORT AND DIFFUSION MODELING
AND ITS APPLICATION
1980

March 1, 1980
(Revised December 29, 1980)

Prepared by:
Maynard E. Smith

For: The Steering Group of the
AMS/EPA Cooperative Agreement

Meteorological Evaluation Services, Inc.

December 29, 1980

TRANSPORT AND DIFFUSION MODELING
AND ITS APPLICATION
1980

Maynard E. Smith

I. SUMMARY

In October, 1979, the American Meteorological Society (AMS) and the Environmental Protection Agency (EPA) entered into a cooperative agreement whereby AMS would provide expertise and assistance to EPA in evaluating air quality modeling[1]. This paper has been prepared at the request of the Steering Group of the AMS/EPA Cooperative Agreement as an overview of current diffusion modeling as it is employed by EPA and other agencies in the United States. Because of the close relation between modeling and air quality regulatory procedures, the paper is as concerned with the use of the models as with their technical quality.

Meteorologists have consistently recognized that no model, however elaborate, can reproduce the full range of complexity of atmospheric motions, and even if it could there would be no hope of measuring those motions completely enough to provide perfect input data. This by no means destroys the value of mathematical modeling, but it should constrain the user to remember, and to heed the fundamental limitations.

Unfortunately, models are being misused in the United States, and it is not generally recognized that these misapplications have arisen largely from the actions of overzealous legislators and regulators, rather than from the choice of the modelers. Confidence in both models and modelers has therefore been eroded, and attention is focused on the weaknesses rather than the strengths of modeling. In fact, we are probably in the favorable position of having excellent modeling capabilities, but we sorely need to provide ourselves with the data to prove it.

Therefore, it is my conviction that the modeling fraternity, via the American Meteorological Society, should take a firm stand in opposition to the misuse of models in regulatory procedures; it is currently unscientific and should be labeled as such. We should take an equally firm stand in favor of obtaining improved data upon which to base modeling developments and applications. The latter would require only modest revamping of the National Weather Service observational network, but a large expansion in field experiments.

Given significant improvements in the supporting data, one may anticipate that improvements in modeling will follow almost automatically. Interest and support for research in model development itself is already adequate, and the necessary computer facilities are a improving faster than any other branch of technology.

II. BACKGROUND

There is no need in this paper for an extensive technical analysis of diffusion modeling, nor for a lengthy history of its development. Recent papers by Turner[2] and DeMarrais[3] have covered these subjects well. Also, as part of this cooperative project, the Steering Group has requested papers on specific aspects of transport and diffusion modeling, each of which will contain a detailed review of one or more of the subjects covered generally in this paper.

My emphasis is therefore on the status of diffusion modeling, stressing the main concepts and assumptions used in various categories of models, their inherent strengths and weaknesses, and the research which might lead to genuine improvements. Since this paper is specifically related to regulatory problems, I have also considered what our society is demanding of diffusion modeling, and what the diffusion models, could they express themselves, should be demanding of us.

Like most other scientific specialities, transport and diffusion modeling has gone through definable phases. The first, from about 1920 to 1947, was a period in which most of the basic theoretical concepts now used in air quality modeling were developed, but during which experimental research was considered a necessary handmaiden to theory. L. F. Richardson and G. I. Taylor were, and continue to be key figures in the development of diffusion modeling, and one need only mention the close relationship between the first use of statistical expressions to represent turbulent diffusion by O. G. Sutton[4], and the field experiments at Porton, England, to be reminded of the attitude of the scientists working in that period. It is certainly to their credit that most of the transport and diffusion models used today use many of the key assumptions which appeared 35 years ago.

The period from 1948 through 1968 might be characterized as the experimental phase. There was an intense desire on the part of the turbulence and diffusion fraternity to understand the physics of the atmospheric processes better, and surprisingly elaborate facilities and experiments were used to further this understanding. The Greenglow, Prairie Grass and the LAPPES programs, as well as long-term studies of plume behavior at TVA and in the United Kingdom are important examples. There were also many studies of the properties of the atmospheric boundary layer, mainly conducted on tower installations at National Laboratories and universities. Improvements in the theoretical treatment of turbulence and diffusion consisted largely of refinements of existing concepts, based on good experimental data.

In the late 1960's, an important change appeared. Gifford[5] in a recent governmental meeting described this change very well. In discussing his concern over blind reliance on elaborate computer models, he said:

> ". . .for 20 years, from 1948 to 1968, approximately, the thrust of the research that went into the atmospheric part of our problem. . .was to find out more about the parameters that govern the phenomena that concern us, and to try to understand the physics better. This governed absolutely everything that was done. The question of computer applicability was a secondary one. The computer was a tool, like the slide rule, but better. And then, all of a sudden in 1968. . .the problem changed by an order of magnitude in distance and in complexity

> . . .(and) we find a completely different approach."

Gifford's point was that recent work in this field has been dominated by computerized modeling based on remarkably little new observational evidence. The literature has been inundated with models, but in many cases the authors have either ignored comparison with field data entirely or settled for comparison with other models, most of which are also unverified.

It is not my purpose to assign responsibility for this situation, but it is rather easy to see how it has come about. Part comes from our general enchantment with computers; they are impressive, they produce reams of seemingly reliable numerical output, and they provide easier and cheaper answers than field experiments. Associated with this trend, and of tremendous importance, is the stipulation appearing in documents such as the Clean Air Act and its Amendments that computer modeling will be the basis for regulatory decisions.

III. DEMANDS ON MODELING CAPABILITIES

Until 1968, transport and diffusion models were treated as guidelines, although widely used in designing new facilities and altering old. For the most part, the users were also researchers, and the limitations in model accuracy were recognized. Also in most cases, the models were applied to individual projects and adjustments in the models were made to suit the specific purpose. These adjustments represented the judgment of the user, but there were no restraints on his right to apply them.

After the Clean Air Act was passed in 1967, and particularly following the enactment of the Amendments of 1970 and 1977, the picture changed abruptly. Diffusion models became a part of the law of the land, and as regulatory procedures developed from that law they became the numerical basis for enormously expensive and important decisions. Precise limits on emissions, and even the acceptability of a new source in a given area, are now regulated on the basis of the numbers provided by an officially-recommended set of diffusion models.

In the nuclear field, diffusion models were also established as the basis for judgment and currently, following the Three-Mile Island Accident, the Nuclear Regulatory Commission is requiring much more elaborate computer modeling and data networks to simulate and monitor future nuclear releases.

Models used in this way are no longer guidelines, they are instruments of national policy.

Along with this fundamental development of governmental policy, simple models have been applied to complex problems. Such phenomena as acid rain, the conversion of SO_2 to sulfates, CO_2 buildup, the mortality and morbidity associated with pollutant concentrations, have all been evaluated without demonstrable improvement of the models or the data on which they are based.

It also seems clear that, as the computer output sheets grow and are disseminated, modeling results are being accepted with less and less regard for their validity. Many are simply treating modeling results as though they were categorically accurate. Two examples serve to illustrate the point. Recently the state of Pennsylvania appealed to EPA for relief from an alleged particulate burden coming from other states upwind of its own borders. The primary basis for this request was a set of rudimentary, speculative model calculations purporting to represent particulate loading derived from sulfates. Similarly, the New York Times, in its magazine section, recently released an article whose title and text strongly suggested that thousands may be dying in the eastern United States because of sulfates generated in the Midwest. The basis for these speculations is modeling which contains a string of tenuous assumptions about the relation between health effects and very low concentrations, as well as transport, diffusion, deposition, depletion and washout.

There is no better term to describe the misuse of modeling in both the governmental and speculative publications than "bad science." The late Cyril Comar published an excellent editorial in Science[6] in which he described "bad science" as a tendency to draw hard conclusions from inadequate data and to base regulatory policies and procedures on such conclusions. I feel sure that, with regard to these two examples, Comar would have suggested that the authors use the models to indicate the range of possible results, respecting the errors inherent in both the data and the models. Had they done so, the modeling would have served a useful purpose in alerting us to possible problems instead of providing a basis for dogmatic conclusions and vitriolic arguments.

IV. MODELS AND MODELING CONCEPTS

A. Plume Rise

The initial behavior of buoyant stack plumes has been investigated more thoroughly than most other atmospheric processes. It has also been reduced to mathematical approximations by a number of researchers, but the studies of Briggs[7,8] stand as the most complete available. Selected portions of Briggs' work have been incorporated into many diffusion modeling systems including those officially recommended by EPA. While more research in this specialized area is desirable, the modeling of plume rise is generally sound and more reliable than other parts of diffusion evaluations.

The behavior of effluents more dense than the ambient air (plume subsidence) has not been studied as extensively, and there is much less experimental data available. Admittedly, releases of dense gas usually occur during accidental spills rather than in orderly routine operations subject to emission control, but the problems which may be created by such releases are by no means trivial.

1. Strengths

The formulas for plume rise associated with momentum and positive buoyancy seem to be based on sound theoretical principles, and they have been reasonably well validated.

2. Weaknesses

a. The most important deficiency is the lack of appropriate input data for use with the formulas. One needs the wind and temperature structure in the layer in which plume rise is occurring, as well as the depth of the convective mixing layer, in order to make proper use of the formulas. Seldom are these quantities measured directly.

b. Combined plume rise-diffusion computer codes often do not take account of the plume mixing induced by the buoyant rise through the ambient air. For very large plumes, this is often an important adjustment, and it is a welcome addition to the most recent EPA codes.

There is also contradiction in the way in which various systems take account of plume rise. In nature, a buoyant plume rises at a gradually decreasing rate, mixing rapidly with the ambient air at first because of its own motion through it, but eventually diffusing at a rate determined by the ambient turbulence as the rise phase comes to an end. Some systems take account of both the transitional rise and the induced mixing; others represent the transitional rise only; still others treat the plume rise as though it occurs directly above the stack top, without induced mixing. The physics of the problem suggests the first of these three alternatives as the most reasonable.

c. There is one aspect of plume rise that has not been treated in any recommended EPA model, and this is the variability in effective plume height arising from variations in the horizontal wind speed over periods of an hour or more. This point was noted by Pasquill in 1974[9], and Egan and Bass[10] estimated the importance of the effect in 1976. It was also emphasized in the recent EPA workshop on instruments and measurements for diffusion studies[11].

d. Experimental data on the behavior of dense gases is still meagre, although the recent NATO-sponsored conference in Rome[12] shows that there is considerable work going on in Europe.

B. Steady-State Gaussian Diffusion Models

Almost all of the transport and diffusion models now used routinely in this country are based on the steady-state Gaussian point-source formulas developed by O. G. Sutton in the 1940's. Many elaborations have been added, such as limits to vertical and horizontal diffusion, and adjustments for transformation, deposition, decay, buildup and depletion. Computers have enabled us to deal with multiple source and receptor configurations and years of continuous input data, but the core of the calculations is still the Gaussian statistical assumption. It is therefore important to review thoroughly the strengths and weaknesses of Gaussian models.

1. Strengths

a. The simple continuous point-source version of this model, when used under the following restrictions, gives good results:

1. continuous, steady pollutant release
2. non-buoyant emission
3. source height $\leq$ 150 m
4. distances $\leq$ 10 km
5. non-transient meteorological conditions

b. Gaussian models are very adaptable. Since they are statistical, they do not depend upon unalterable physical assumptions, and they are easily adjusted to match observational data. For example, the long-range predictions of the steady-state models might be improved significantly by altering the expressions used to describe the vertical and horizontal diffusion parameters, σz and σy, if we simply knew more about their variability with distance. Similarly, it would not be difficult to introduce better adjustments for terrain height, limits to diffusion, decay, deposition, etc. if the variables governing such phenomena were known more accurately.

Note that this flexibility maybe lacking in a model which is based on rigorous physical reasoning. When such a model performs badly, it is usually necessary to review the entire set of assumptions on which it rests.

c. In general, this family of models is easy to use; they can applied to rather lengthy sets of data at acceptable cost.

d. Not the least of the advantages of the steady-state Gaussian models is familiarity. Most individuals seriously involved in diffusion studies have had some training in their use, and can apply them in an orderly fashion.

2. Weaknesses

a. The most serious shortcoming of these models is their inflexibility with respect to variable flow trajectories and changes in meteorological conditions. It may be unfair to list this quality as a shortcoming because these models were never designed to handle such conditions, but since they are being used to represent variable meteorology, the deficiency must be acknowldeged.

This weakness is most serious in two very different applications of the models, long-range transport and accidental releases.

In the former, wind direction, wind speed and stability inevitably change with both time and distance; in the latter, one needs a real time and space representation of the location and magnitude of concentrations from a source which is highly variable. In either application one should anticipate poor results with steady-state models.

Attempts have been made to adjust steady-state models for changes in certain parameters with time and distance, and at least one hybrid model accommodates variable wind trajectories, but there is little evidence that these modifications give realistic results.

b. The existing Gaussian models, as they are typically used today, are incapable of representing wind shear. Wind speed and direction shear in the vertical can probably be ignored in short-range applications under neutral and unstable conditions, but it becomes important, perhaps even dominant, at very great distances.

c. Gaussian models do not represent diffusion under stable conditions very well. In these situations, there is so little vertical motion that diffusion in this dimension depends largely on processes associated with the release of the pollutant, while lateral diffusion is dependent almost entirely upon wind direction shear and plume meander[13]. It is doubtful that either the vertical or horizontal distributions are Gaussian, and a different model would probably be more appropriate.

d. Steady-state models cannot describe calm, or near-calm conditions. By definition the models represent a continuous air flow, and there is nothing within the framework of the model that permits representation of these common conditions.

e. Steady-state models are probably at their worst in extremely rough terrain, where complicated three-dimensional flow is usually present. Even if the flow pattern were known in full detail, it is difficult to visualize any way in which such a model could handle the complexity.

A particular aspect of this deficiency, which has received much attention but no accepted solution, is the so-called "terrain correction." This term means the adjustment in the concentration prediction made to account merely for the fact that the underlying surface varies significantly in height compared to the base elevation of the source. It would appear that a proper correction varies between a "full" adjustment, in which one subtracts the entire difference between source and receptor elevation from the effective plume height, and some fraction of this value, depending upon the variability and slope of the underlying surface.

When the terrain height significantly exceeds that of the effective plume height, it is difficult to suggest a solution that would be both reasonable and simple.

It seems clear the remedying these deficiencies requires a variable trajectory model capable of depicting changes in the three-dimensional flow patterns.

C. Stability Classes and Dispersion Parameters

Since Gaussian plume models represent the "state of the art," at least in present applications, it is appropriate to review two key subjects associated with these models, the methods of defining atmospheric stability and the specifications of the dispersion parameters associated with them. These two subjects are of course equally important to other families of models, but their close association with Gaussian models suggests reviewing them at this point.

1. Stability Classes

The structure of turbulence within the planetary boundary layer is dependent on a complex interaction involving the air flow, surface roughness, radiative processes and major terrain features. Clearly, we need some means of representing the influence of these factors on turbulence. In the past, both for simplicity and because of inadequate knowledge of conditions in the boundary layer, we have tended to focus on one or another simple indicator to derive stability estimates. The most common of these are the STAR system, which is a method of using standard surface weather observations to estimate stability, the Delta-T technique in which temperature difference over a selected height interval is the indicator, and schemes which use the fluctuations of a horizontal wind as the indicator. Only in research studies have more elaborate methods been used to represent atmospheric stability and turbulence.

As might be expected with simplified yardsticks, none of the these three systems does an adequate job. There have been many comparisons among the systems, showing that each provides remarkably different distributions of stability classes. The situation may be even worse than the literature indicates because many of the important pollutant sources are far above the near-surface layer in which these indicators are usually measured.

2. Dispersion Parameters

The most common specification of the diffusion parameters, σy and σz, is the Pasquill-Gifford system[14].

These curves are based on experimental data taken in short periods of time over smooth terrain at low elevations close to the source, and there is considerable evidence that they are not appropriate for other applications, particularly for very high, buoyant sources over longer periods of time.

This particular problem serves as a good example of the difficult interplay between the scientific and regulatory viewpoints. The regulators naturally want as simple a system as possible, to be applied consistently to a wide variety of situations. The Pasquill-Gifford sigma parameters provide such a system. The scientists, however, recognize that these expressions for the sigma values represent some situations much better than others, being quite good for local effects from low-level sources and demonstrably poor for buoyant, elevated plumes. This controversy has focused on the application of the sigma values for the most unstable category, Stability A, to the determination of emission limitations for large utility and industrial sources, wherein a modest inaccuracy can have mutli-million dollar implications. Failure to resolve issues like this outside the courts has been detrimenteal to both scientists and regulators.

There is one important deficiency which has been stressed by Irwin[15]; neither stability regimes nor dispersion parameters are segregated into discrete classes by nature. Nature favors a continuum of stability situations and the division into classes is a man-made convenience. Employing this convenience has the effect of fixing the impact of a particular source into rigid patterns, and if a site has a high frequency of a given stability class coupled with a narrow range of wind speed, a distortion of the concentration distribution will be brought about by the classification system itself.

The most recent EPA workshop having to do with instrumentation and measurements for diffusion modeling[11] highlighted these differences, and pointed the way toward a better set of estimates in the future. Whatever models are used, these two factors must be better defined.

D. Box and Pie-Shaped Models

Although it may seem unsophisticated to include the simplest of models in this overview, probably there is merit in doing so. Models of this type usually assume uniform mixing of the pollutant within a geometrically defined volume, and such models have a number of advantages. They are extremely simple and inexpensive to use and the assumptions involved are both perfectly clear and very easy to alter. In addition, they can be used to set reasonable limits for comparison with more elaborate models. Not the least of these advantages is the high probability that, at present, such models give predictions just as accurate as much more sophisticated systems when applied to long-range transport or long periods of time.

E. Time and Space-Dependent Models

Within the past ten years many more sophisticated models have appeared, most of which are designed to depict diffusion and physiochemical transformation in a three-dimensional wind and turbulence field which can vary with time as well as space. These models, involving many different concepts, are described as advection-diffuison models, variable trajectory models, particle-in-cell models, and grid models.

Whatever their internal workings and descriptions, they offer potential benefits over the steady-state Gaussian models, particularly in their ability to follow the actual trajectories of plumes and clouds, and to represent more accurately the time history and transformations of non-conservative pollutants.

Unfortunately, most of these models share another attribute, namely an almost total lack of validation with actual field data. The typical publication describing one of these newer models stresses the elegance and validity of the mathematics and assumptions involved, but actual performance is usually ignored altogether or assessed only by comparison with the performance of other models.

While some of these models will undoubtedly become part of our arsenal of transport and diffusion techniques in the future, it is difficult to make intelligent assessments of their applicability or accuracy at present. The best that one can do is assess the physical assumptions and mathematics of a given system, but no model available is devoid of simplifying assumptions and none can really be assessed without excellent field data for comparison.

With respect to validation data, it is important to note that there are some relatively untapped sources of information; sources which can be used with a bit of effort and some administrative cooperation. The Lappes plume data and the long-range transport data from the Savannah River Laboratory are pertinent examples.

F. Fluid Models

Transport and diffusion modeling is not customarily construed as including fluid models, at least when discussed by meteorologists. However, there are two reasons for doing so in this overview paper. First, proposed EPA regulations specify fluid modeling as an acceptable procedure for the solution of certain types of GEP stack height problems, and second, fluid modeling is preferred for the solution of certain complex problems.

When moderate to strong winds prevail in rough terrain, the flow patterns close to the source are determined largely by aerodynamic forces, and in such cases fluid modeling is far more likely to depict the true distribution of concentrations than any calculational model.

It also seems probable that fluid models may prove valuable in representing atmospheric behavior under stable conditions in rough terrain. Plans are already under way to compare fluid modeling results with field experience under such conditions[16].

The advantages of fluid modeling are obvious. The cost of examining a problem in a tunnel is very much smaller than that of a full-scale field study, and it is much easier to control test conditions.

G. Recommended EPA Models

EPA has adopted a series of diffusion models to represent various source-receptor configurations, input data and concentration output formats. These models are available from NTIS under the general designation of UNAMAP models, and with two exceptions, they are fully described elsewhere[17].

All of these models are based on the Gaussian assumption, and they therefore partake of the strengths and weaknesses of other Gaussian models. There are, however, a number of problems specific to the models EPA has recommended.

1. Inconsistencies

There are two types of inconsistency involved in current applications of transport and diffusion modeling. The first derives from different treatment of identical technical problems in EPA-recommended models, giving rise to directly conflicting results when two models are applied to the same problem. Clearly this situation should be avoided wherever possible. These inconsistencies have included, for example: different exponents for wind profile adjustments, different reference heights for wind estimates, different methods for terrain correction, and different treatment of plume rise. These inconsistencies probably arose because the models were developed by different individuals or groups, and as long as the models were actually used as guidelines instead of a means of specifying emission limitations, they were trivial. However, the inconsistencies became serious when the models entered formal regulatory and legal proceedings.

There is no doubt that EPA is aware of this problem, and action is being taken to eliminate technical inconsistencies among the existing models. Presumably, new models which are now being released, such as MPTER and ISC, will be consistent with each other.

The second type of inconsistency is one in which models are applied differently in different EPA regions and offices. For example, one region might use a simple Gaussian model on a problem that another would treat with a complex terrain code. This type of inconsistency is administratively chaotic, but scientifically desirable because without some freedom, modeling will truly become a completely mindless exercise.

2. Scope

EPA has not recommended a model capable of handling variable plume trajectories or representing real-time pollutant distributions. Testing and validation of such a model presents formidable problems, but the need exists.

Furthermore, one should have guidance in modeling accidental releases. EPA has devoted most of its energies to continuous releases, but intelligent management of accidents is important also. It should be remembered that any model designed to represent accidents should be capable of dealing with dense gases such as chlorine and liquified natural gas.

EPA is cognizant of the need to expand the scope of its predictive capability, and steps are being taken to remedy the more obvious gaps. MPTER and ISC should improve the multiple-source treatment, but as yet there seems to be no candidate for a variable trajectory model.

3. Rough Terrain

A very serious defect is the lack of a good model for use in rough terrain. The VALLEY model is at best a rough approximation and claims for its accuracy, or the lack of it, are conflicting.

EPA is aware of this problem, and the agency is embarking on a major study of diffusion in rough terrain[16]. It is not likely that a single study will clearly identify a particular model as the best for complex terrain, but it is a step in the right direction.

4. Validation

By far the most serious shortcoming of all EPA models is lack of validation. There has been no concerted attempt to assess the accuracy of the UNAMAP series which now form an integral part of national regulatory policy. Admittedly, there are relatively few sets of data ideally suited to the validation procedure, but this issue cannot be ignored. The costs associated with air quality control are too high to be based on unvalidated techniques.

Improved validation studies would undoubtedly reduce the controversy over choices of input data, stability estimates and sigma curves, to name a few salient controversial topics.

Recently the AMS/EPA Cooperative Agreement Group conducted a workshop on evaluating model performance[21], and one hopes that these efforts will be developed into an active program by EPA.

5. **Workshops**

One of the more important recent activities in the modeling field, or directly affecting it, has been the use of workshops to review specific topics.

Workshops have the advantage of concentrating the attention of a small number of specialists on a particular problem, and the requirement of completing a first draft of a report during the workshop sessions crystallizes the results.

Within the past three years there have been a number of workshops of direct importance in transport and diffusion modeling, one on the EPA Modeling Guideline[18], two on stability estimates and dispersion parameters[19], [20] one on complex terrain[16], and one on instrumentation and measurements for air quality modeling[11], and one on model performance and validation[21].

EPA deserves commendation for its active support of these workshops. It has sponsored a number of them, participated in all, and translated the results into positve action.

V. QUALITY OF INPUT PARAMETERS

The representativeness and accuracy of the transport and diffusion models is being hotly debated throughout the country, but one seldom hears a corresponding debate about the accuracy of the input information. Perhaps the simplest way to illustrate the importance of the input data is to imagine that we had at our disposal a perfect transport and diffusion model. Such a model could handle any degree of complexity in the atmospheric physics and chemistry without significant error. Clearly, the accuracy of the model output would then depend solely upon the accuracy and completeness of the input information. This hypothetical situation becomes meaningful if one considers specific problems which are of direct concern today.

In the following sections, I have estimated the range of accuracy (or inaccuracy) to be expected when the hypothetically perfect model is applied to the prediction of typical limiting concentrations in the U.S. regulatory system, namely to a few values at the extreme upper end of the concentration distribution. The accuracy would undoubtedly improve if the application were made to less capricious statistical quantities, the 95th percentile for example.

It is also true that accuracy would improve if one were considering a receptor in some unspecified location rather than a specific one. In other words, a calculated maximum concentration of 10 ug/m^3 at a given location might miss by a large factor, but agree better with the actual overall maximum measured elsewhere.

In any event, the point of this section is to illustrate the paramount importance of new field experiments and more appropriate input data, regardless of the quality of our modeling capabilities.

A. Low Continuous Source - Local Concentrations

The first is the least complicated; a low-level continuous source, for which we need to know only the local distribution of pollutant concentrations. Such a source closely approximates those in the research studies upon which most of our current knowledge of diffusion is based. Let us assume that the site is instrumented reasonably well, although most sites are not. Table 1 shows in condensed form how we would acquire the input data for the modeling, the quality of the previous research work on which the data are based (if applicable), and finally the probable accuracy of the particular input parameter. No references are given for these estimates of accuracy, because one can vary them over a broad range without invalidating the argument.

Emission rates from these low sources are usually estimated and can be relied upon to no better than ±30%, (an actual/measured ratio of .7-1.3). We have reason to be more confident of the wind and stability data, which we assume to be measured over the proper height interval on an onsite tower. Estimates of σy should be quite good if based on the horizontal wind fluctuations, but σz will probably be somewhat less accurate because this parameter has almost never been properly measured, even in the best of the research studies. Therefore, we continue to be somewhat unsure of the relation between plume σ_z and the corresponding fluctuations of the vertical wind. The effective plume height would be derived from a fixed-stack height plus a modest plume rise, and should be fairly accurate.

TABLE 1

REQUIRED PARAMETERS

SOURCE: LOW-LEVEL CONTINUOUS EMISSION, FLAT TERRAIN

REQUIRED OUTPUT: LOCAL MAXIMUM AND AVERAGE CONCENTRATIONS

	DERIVED FROM:				
PARAMETER	MEASUREMENT	EXTRAPOLATION	ESTIMATE	QUALITY OF FOREGOING RESEARCH	RELIABILITY RATIO (A/M)*
EMISSION RATE			●	NA	.7 - 1.3
WIND DIRECTION	●			NA	.9 - 1.1
WIND SPEED	●			NA	.9 - 1.1
STABILITY CLASS	●			GOOD	CORRECT
σ_y	● (σ_θ)			GOOD	.8 - 1.2
σ_z	● (σ_ϕ)		●	FAIR	.6 - 1.4
EFFECTIVE SOURCE HEIGHT	●	and ⟶	●	GOOD	.85 - 1.15

NA - Not applicable or not available

*A/M - Actual value / Measured or estimated value

However, even with this well-instrumented, simple source in simple terrain, the range of accuracy of the input information is broad, and it would not be at all surprising to find a given calculated concentration differing from measured values by ±50%.

B. Tall-Stack Source - Local Concentrations

A similar list for a typical tall-stack source shows a significant deterioration of the quality of the input data. With the exception of the pollutant emission, which might be more accurately known, the parameters would have much wider ranges of uncertainty. The wind data would be extrapolated from measurements at lower elevations, and the stability class determination would have to be listed as an estimate, whether based on a surface observation system like STAR or on any actual stability determination from a lower elevation. The σ_y and σ_z estimates would suffer correspondingly, not only because of the uncertainty in the stability designation but also because we have less research data on high-level plumes. Finally, we must be concerned with a new parameter, the depth of the convective mixing layer. This is most important for tall-stack sources, but it is crudely interpolated from twice-daily measurements, usually made at a site remote from the source. Continuous data, such as might be available from remote sounding techniques, would be preferable.

With these much larger uncertainties, our perfect model would be doing well indeed if a prediction matched the concurrent observation within a factor of two or three.

TABLE 2

REQUIRED PARAMETERS

SOURCE: HIGH-LEVEL, CONTINUOUS EMISSION, SIMPLE TERRAIN

REQUIRED OUTPUT: LOCAL MAXIMUM AND AVERAGE CONCENTRATIONS

PARAMETER	DERIVED FROM: MEASUREMENT	EXTRAPOLATION	ESTIMATE	QUALITY OF FOREGOING RESEARCH	RELIABILITY RATIO (A/M)*
EMISSION RATE	■		or ■	NA	.8 - 1.2
WIND DIRECTION	■ and	■		FAIR	.7 - 1.3
WIND SPEED	■ and	■		FAIR	.7 - 1.3
STABILITY CLASS			■	FAIR	TWO CLASS INTERVALS
σ_Y			■	FAIR	.5 - 1.5
σ_Z			■	POOR	.30 - 3.0
EFFECTIVE SOURCE HEIGHT	■ and	■		GOOD	.75 - 1.25
HEIGHT OF MIXING LAYER			■	FAIR	.5 - 1.5

NA - Not applicable or not available

*A/M - Actual value / Measured or estimated value

C. Tall-Stack Source - Long-Range Concentrations

Table 3 extends the tall-stack problem to the realm of long-range transport and includes additional factors such as chemical transformation and removal processes. Tracing the origins of sulfates and acid rain depends upon these and other factors. Here all of the uncertainties of the simple, tall-stack dispersion are compounded by a host of other approximations, most of which are little better than rough estimates. The non-meteorologist, for example, may not be aware of the degree of uncertainty about the wind and stability field over a large area. It is true that the United States is studded with a dense network of weather stations, but few of these obtain data at heights more than 10 meters above the surface. Even those which do fail to measure the parameters necessary for transport and diffusion models.

TABLE 3

REQUIRED PARAMETERS

SOURCE: HIGH-LEVEL, CONTINUOUS EMISSION,

REQUIRED OUTPUT: DISTANT CONCENTRATIONS, CONVERSION PRODUCTS

	DERIVED FROM:				
PARAMETER	MEASUREMENT	EXTRAPOLATION	ESTIMATE	QUALITY OF FOREGOING RESEARCH	RELIABILITY RATIO (A/M)*
	(ALL OF TABLE 2) PLUS:				
TRANSPORT WIND FIELD	■	and → ■		FAIR	.5 - 1.5
WIND SHEAR			■	POOR	.25 - 4.0
STABILITY FIELD	■	and → ■	or → ■	NA	TWO CLASS INTERVALS
σ_Y			■	POOR	.5 - 1.5
σ_Z			■	POOR	.30 - 3.0
POLLUTANT CONVERSION RATES			■	FAIR	.1 - 10.0
DEPOSITION RATES			■	FAIR	.3 3.0
RAIN FIELD	■	and → ■	or → ■	NA	.25 - 4.0
WASHOUT RATES			■	FAIR	.25 - 4.0

NA - Not applicable or not available

*A/M - Actual value / Measured or estimated value

The accuracy which our excellent model could achieve under these circumstances might often be wrong by a factor of 5 to 10.

D. Low-Accidental Source

The final example of the series is the accidental release from a low-level source; a chemical explosion or a nuclear accident would fall in this category. In this instance, the most serious deficiency is usually non-meteorological; the source term can seldom be specified to better than a factor of ten. With such uncertainty the defects in the other measurements or estimates tend to become insignificant; however, the meteorological parameters taken at only one or two locations are also less useful than in the case of a continuous point source, because one usually is interested in defining the time and space distribution of the concentrations as well as their overall levels.

TABLE 4

REQUIRED PARAMETERS

SOURCE: LOW-LEVEL ACCIDENTAL RELEASE

REQUIRED OUTPUT: SPACE AND TIME-DEPENDENT CONCENTRATION AND DEPOSITION PATTERNS

	DERIVED FROM:				
PARAMETER	MEASUREMENT	EXTRAPOLATION	ESTIMATE	QUALITY OF FOREGOING RESEARCH	RELIABILITY RATIO (A/M)*
Q			●	NA	.1 - 10.0
WIND FIELD	●	and ——→	●	NA	.5 - 1.5
STABILITY FIELD	●	and ——→	●	NA	ONE CLASS INTERVAL
σ_Y			●	NA	.5 - 1.5
σ_Z			●	NA	.2 - 5.0
EFFECTIVE SOURCE HEIGHT			●	FAIR	.5 - 1.5
RADIOACTIVE DECAY RATES			●	EXCELLENT	1.0**
CHEMICAL TRANSFORMATION RATES			●	POOR	.30 - 3.0
DEPOSITION RATES			●	FAIR	.1 - 5.0

NA - Not applicable or not available

*A/M - Actual value / Measured or estimated value

** - Depends on Composition of Q

The best one can hope for in a situation like this is a prediction of the approximate area covered by the effluent, and the concentration within it could probably be estimated to no better than an order of magnitude.

VI. DEMANDS OF THE MODELS AND MODELERS

If it is really necessary to make far-reaching national decisions with respect to air quality management immediately, it is equally important to assure ourselves that these decisions are based on the best possible information, and that the limitations of that information are clearly understood. In my view, the meteorological fraternity should adopt a well-publicized position with respect to the following points:

A. Misuse of Air Quality Models

It is obvious that no air quality modeling study can be extremely precise, now or at any time in the foreseeable future. The lack of precision derives from the fundamental complexity of the atmospheric processes and our inability to measure them completely. Transport and diffusion models should therefore never be used precisely either to specify emission limitations or to estimate airborne concentrations.

If one accepts this limitation, it is fair to ask how models should be used by the regulatory agencies; should they be a part of the decision-making process at all? The answer is clearly "yes." Man owes it to himself to gauge the consequences of his actions, and modeling is a necessary part of this process.

Diffusion models are very valuable now as guidelines in estimating the local effects of emissions on air quality, and the results can be quite reliable when the modeling is keyed to suitable meteorological data, air quality data and reasonable specification of the limiting air quality objectives. For example, it is no accident that most large SO_2 emitters designed and built since 1960 produce local ground-level concentrations meeting our present Federal standards, because most such sources were considered from the standpoint of keeping hourly concentrations below approximately 0.5 ppm SO_2.

This earlier decision-making, however, might have included choosing a 600-ft. stack over a 550-ft. stack to keep emissions ranging from 3.5 to 4.5% sulfur from causing local hourly concentrations from exceeding 0.5 ppm very often. Now our regulations transform this same process into a decision involving a 587-ft. stack (perhaps chosen after a 2-year wind tunnel study) 4.16% S, to avoid exceeding 0.50 ppm more than once per year, with all the legal wrangling that goes with it.

Even in the field of long-range transport and transformation, the models provide useful guidelines. However, they are necessarily far less accurate in this type of application, and they certainly should not serve as the basis for interstate lawsuits.

As we begin to acquire more helpful data, the models, or at least their application, will improve and they will become more reliable guidelines. Perhaps by 1990 they will become 12-inch rulers instead of yardsticks, but they will not be micrometers.

B. Misuse of Fluid Modeling

Similarly, fluid modeling facilities will be grossly misused if proposed regulations for determining Good Engineering Practice (GEP) stack heights are implemented in their present form. Such regulations promise to clog the existing wind and water tunnel facilities with a host of elaborate aerodynamic projects designed to produce seemingly precise results for setting emission limitations. The fluid models cannot represent the atmosphere with the accuracy demanded of them, and they should never be used to do so.

On the other hand, fluid modeling has not been explored as fully as it should be with respect to representing flow in complex terrain, particularly under stable or vertically-limited atmospheric conditions. The complex-terrain problems will probably prove to be the most difficult to model theoretically, and fluid models should be used to the fullest extent possible to help solve them.

C. Need for Better Meteorological Data

The data obtained by the National Weather Service is geared almost exclusively to aviation and general forecasting requirements. Air quality is now also a major national objective, and the observational network should be altered to reflect this fact.

The most important reasons for involving our national system are simplicity and cost. The data are needed over the entire country, and the NWS and the National Climatic Center already have the network, data-processing, archiving and retrieval facilities required to manage the information. Any other solution would probably be far more difficult to implement, and the results less useful and consistent. In the interest of consistency, it would also be helpful if EPA and NRC could coordinate their requirement before such a program is started.

There is no need in this overview paper to list precisely the recommended changes in the NWS system. These have been carefully defined in the recent EPA-sponsored workshop on instrumentation for air quality evaluations[9], and they include better definition of the important "surface" parameters, as well as improvement in measurements aloft. The recommendations for the latter carefully avoid instruments or procedures that would be inimical to aircraft operations.

The requirements for more suitable data must not stop with the NWS; the industrial and research communities must also be involved. Major new sources should be required to obtain more suitable onsite data, presumably processed to fit directly into the national archives. Such data must include better estimates of the wind and turbulence at source height, as well as the overall depth of the convective mixing layer. Fulfulling the requirement for remote measurements should be a continuing objective of the instrument research people.

D. Need for Field and Laboratory Studies

The best possible estimates of the environmental effects of pollutant emissions on regional and national scales depend upon proper data. One would hardly extrapolate the parameters suitable for a two-story frame house to a structure the size of the World Trade Center, but that is what the air quality modeler is now being asked to do.

Large-scale transport and dispersion research studies are expensive, in the multi-million dollar range, but the consequences of inappropriate emission regulations are in the multi-billion dollar range.

We already have several excellent research projects of this type either planned or underway. Projects like SURE, the EPRI three-phase power plant study, and the EPA rough terrain project seem well-conceived and they should provide insight into some of the important problems.

However, we should not deceive ourselves into believing that completion of three or four major projects will settle all controversy and advance the quality of modeling enormously. One cannot possibly foresee all of the instrumental and logistic difficulties that will plague a major project, nor can one prejudge correctly the relative importance of the various parts of the work. The results will always be to some degree incomplete and inadequate, suggesting the need for additional studies.

There is a need, therefore, for less ambitious projects, in which limited, well-defined objectives are studied. Such studies would illuminate specific factors more clearly, and serve as a check on conclusions reached in the larger-scale projects. These lesser projects may range from laboratory research to large field investigations, and they should include, at the very least, the following subjects:

1. Formation of atmospheric sulfates and nitrates.

2. Removal of pollutant gases and particles by dry and wet processes.

3. Definition of the characteristics of acid precipitation, as a function of time, space and ambient conditions.

4. Influence of wind shear on dispersion.

5. Vertical and horizontal diffusion from 10 to 100 km.

6. Intensive studies of local and regional maximum pollutant concentrations.

7. Effect of moderate terrain fluctuations on pollutant concentrations.

8. Structure of the planetary boundary layer.

9. Synoptic studies of pollutant distibutions.

10. Effects of small particulates as well as CO_2 in the global heat balance. One may tend toward cooling while the other favors heating.

11. Validation and performance studies.

E. Peer Group Review

Comar[6], in his editorial on "bad science," deplored the fact that many publications are not subjected to the scientific peer review which was routine two decades ago. Rather, they are directly released from governmental agencies, major laboratories, universities and consulting agencies without external scientific review.

Certainly this is a major weakness in the present EPA procedures. Guidelines and publications may appear to have undergone such review, but there is actually very little evidence that the scientific community is having significant influence in this area.

F. Modeling Research

One might anticipate a strong recommendation for support of additional research in modeling, and perhaps for related computer development that might make the modeling more effective and practical. Neither seems warranted, in my opinion. One can be confident that new sources of field data will be devoured by the modelers as quickly as they become available. If some of the newer modeling concepts are sound, the data will prove them so; if there are serious gaps, they will be recognized and accommodated. For their part, the computer specialists certainly need no spur from the meteorologists to continue their remarkable development pace.

The paramount need, it would seem, is data, not just the development of more elaborate models.

REFERENCES

1. _____Air Quality Modeling, Notice, AMS Bulletin, V60, #10, (1979).

2. Turner D. B.: Atmospheric Dispersion Modeling, A Critical Review, Jour. APCA, V29, #5, (1979).

3. De Marrais G. A.: A History of Air Pollution Meteorology Through 1969, NOAA Tech. Memo, ERL-ARL-74, (1979).

4. Sutton O. G.: The Theoretical Distribution of Airborne Pollution From Factory Chimneys, Quart. Jour., Royal Met. Soc., 73, (1947).

5. Gifford F. A.: Transcript, Subcommittee on Site Evaluation Meeting. ACRS. Nuc. Reg. Comm. (Ace-Fed. Reporters), Dec. (1979), pp. 103-104.

6. Comar C.: Bad Science and Social Penalties, Science V200, #4347, (1978).

7. Briggs G. A.: Plume Rise, USAEC Critical Review Series, T10-25075, NTIS, (1969).

8. Briggs G. A.: Plume Rise Predictions, Lectures on Air Pollution and Environmental Impact Analysis, Amer. Met. Soc., (1975)

9. Pasquill, F.: Atmospheric Diffusion, Ellis Horwood, Ltd., Chichester, Eng., Wiley & Sons, (1974).

10. Egan, B. A. and Bass, A.: Air Quality Modeling of Effluent Plumes in Rough Terrain, 3rd Symp. on Atm. Turb. & Diff., Raleigh, N.C. (1976).

11. Environmental Research & Technology, Inc.: On-Site Meteorological Instrumentation Requirements to Characterize Diffusion from Point Sources, Report on Workshop, Raleigh, N.C. Jan. 15-19, (1980) (in press).

12. NATO, CCMS: Proceedings, 10th International Technical Meeting on Air Pollution Modeling and its Applications, NATO, Rome, Oct., (1979).

13. Sagendorf J. F. and Dickson, C. R.: Diffusion Under Low Wind Speed, Inversion Conditions, NOAA Tech. Memo. 52, ERL-ARL, (1975).

14. Gifford F. A.: Uses of Routine Meteorological Observations for Estimating Atmospheric Dispersion, Nuclear Safety V2, #4, (1961).

REFERENCES
(Cont.)

15. Irwin J. S.: Estimating Plume Dispersion - A recommended Generalized Scheme, Proc., 4th Symp. on Turb., Diff. & Air Poll., AMS, Jan., (1979).

16. U.S. EPA: Workshop on Atmospheric Dispersion Models in Complex Terrain, EPA-600/9-79-041, (1979).

17. U.S. EPA: Guideline on Air Quality Models, EPA-450/2-78-027 (1978).

18. Argonne National Lab.: Report to the U.S. EPA of the Specialists' Conference on the EPA Modeling Guideline, ANL, (1977).

19. Hanna S. R., Briggs, G. A., Deardorff, J., Egan, B. A., Gifford, F. A., and Pasquill, F.: AMS Workshop on Stability Classification Schemes and Sigma Curves - Summary of Recommendations, Bulletin, AMS V58, #12, (1977).

20. Randerson D.: Review Panel on Sigma Computations, Bulletin AMS V60, #6, (1979).

21. Fox, D. G., (Working Chairman): Judging Air Quality Model Performance, Amer. Met. Soc. (in press).

Appendix K

Judging Air Quality Model Performance

A Summary of the AMS Workshop on Dispersion Model Performance, Woods Hole, Mass., 8-11 September 1980

Douglas G. Fox[1]
Workshop Chairman

Abstract

Under the direction of the AMS Steering Committee for the EPA Cooperative Agreement on air quality modeling, a small group of scientists was convened to review and recommend procedures to evaluate the performance of air quality models. Particular attention was paid to the operational use of models in regulatory settings.

A number of general recommendations resulted from the workshop. The usefulness of models to aid decision makers in air quality management was reiterated. Concerns about the absolute, rather than the statistical, nature of air quality standards were raised, with a recommendation to reformulate standards accordingly. Model performance evaluation was suggested on the basis of differences between observed and predicted concentrations. The bias (average), the variance (noise), and the gross variability (gross error) of such differences were the recommended performance measures. In addition, correlation measures calculated in time, space, and jointly between observation and predictions were recommended. These measures and suggestions on how to apply them are outlined. Some qualification on the use of these measures for application to point sources using limited data sets was made.

The workshop did not recommend any specific numerical standards for acceptable model performance; there was insufficient information to do this. Rather, considerations for evaluating model performance using statistically constructed confidence intervals and comparison against EPA's currently used models were suggested.

Finally, areas in need of further research were identified. These include development and description of performance measures; application, especially to point source models; analysis of meteorological and air quality data, both existing and newly developed; and evaluation of diffusion modeling techniques.

1. Introduction

Atmospheric dispersion models are used in support of laws and regulations designed to protect the nation's air quality. These models provide a scientifically reasonable connection between a source and a receptor of air pollutants. Courts have interpreted them to be legally binding mechanisms for establishing acceptable levels of emissions control from sources; therefore, models must be correct and must be correctly applied. To meet this need, the U.S. Environmental Protection Agency (EPA) has entered into a cooperative agreement with the American Meteorological Society (AMS) to help develop and apply atmospheric dispersion models.

To evaluate the performance of models, AMS convened a group of professionals on the subject, in September 1980, to discuss current practices in model evaluation, to recommend model performance evaluation measures and methods, and, if possible, to set standards for model performance. An additional task was to identify specific areas needing further research.

The workshop participants are responsible for the content of this report. The participants were: *Douglas G. Fox, Chairman, U.S. Department of Agriculture (USDA) Forest Service, Ft. Collins, Colo.; Robert Bornstein, San Jose State University, San Jose, Calif.; Norman Bowne, TRC Environmental Consultants, Wethersfield, Conn.; Robin L. Dennis, National Center for Atmospheric Research (NCAR), Boulder, Colo.; *Bruce Egan, Environmental Research and Technology, Inc., Concord, Mass.; *Steven Hanna, National Oceanic and Atmospheric Administration (NOAA), Oak Ridge, Tenn.; Glenn Hilst, Electric Power Research Institute (EPRI), Palo Alto, Calif.; Stuart Hunter, Princeton University, Princeton, N.J.; Michael Mills, Teknekron Research, Inc., Waltham, Mass.; Larry Niemeyer, EPA, Research Triangle Park, N.C.; Hans Panofsky, Pennsylvania State University, University Park, Pa.; *Darryl Randerson, NOAA, Las Vegas, Nev.; Philip Roth, Systems Applications, Inc., San Rafael, Calif.; Ronald Ruff, SRI International, Menlo Park, Calif.; Lloyd Schulman, Environmental Research and Technology, Inc., Concord, Mass.; Jack Shreffler, EPA, Research Triangle Park, N.C.; Herschell H. Slater, Consultant, Chapel Hill, N.C.; Joseph Tikvart, EPA, Research Triangle Park, N.C.; A. Venkatram, Ontario Ministry of the Environment, Toronto, Ontario, Canada; Jeffrey C. Weil, Martin Marietta Corp., Baltimore, Md.; and *Fred D. White, Consultant, Arlington, Va.

Members of the AMS Steering Committee for the EPA cooperative agreement (denoted by *), especially Fred White, Chairman of the Steering Committee, provided special assistance in conducting the workshop.

2. Regulatory use of models

An air quality simulation model, as defined at the workshop, is a quantitative method, based on physical principles, for estimating pollutant concentrations in a specific area and period of time. These concentrations vary with source distribution, and meteorological and geophysical conditions. Air quality modeling also involves the various types of input data needed for the model to operate and the scope of the model outputs, such as where and over how long an averaging time the concentration estimate is required.

The primary effects of air pollution result from the ambient air concentrations of pollutants. To estimate ambient air concentrations, the atmospheric pathway from the source of emissions to any locations of concern—a very complex and sometimes seemingly random path—must be considered.

Air quality standards are written in terms of concentrations. National ambient air quality standards are set at a level determined to be appropriate to protect public health and well-being. Models are used for planning and assessing the attainment of these standards.

Models are used for three general types of application. One is to ensure maintenance of and compliance with ambient standards. Generally, this involves considering the interactions among a large number of pollutant sources in order to evaluate changes in concentration that would result from altering emissions from a subset of the sources. The second general process is the "permitting" of new sources, requiring the calculation of the potential concentration contribution from construction or alteration of each new source. The third application of models is a variant of the permit requirement more generally applied to the environmental review process. Air quality simulation models are seeing increased use as a primary component in evaluating effects of air pollution on ourselves and our environment.

---

[1] Rocky Mountain Forest and Range Experiment Station, U.S. Forest Service, 240 W. Prospect, Ft. Collins, Colo. 80526.

Because the models are applied mostly in a formal regulatory setting, often by legal and administrative personnel, they sometimes become elevated to a status above that of a tool for scientists to estimate the consequences of actions. Model predictions, in many cases, may be used to resolve significant economic and social issues (Fox and Fairobent, 1981).

Because of the significance of model predictions and the complexity of the problem, EPA has sought to standardize air quality simulation models and their application; it has developed a draft modeling guideline for wide technical review. This review included a specialists' conference, at which a subgroup discussing validation and calibration of models recommended that "models should meet certain standards of performance if they are to be endorsed in the Guideline for general use" (Roberts, 1977).

EPA was required by the 1977 Amendments to the Clean Air Act to ". . . specify with reasonable particularity each air quality model or models to be used under specified sets of conditions. . ." in the "permitting" process. EPA also was directed to conduct a modeling conference (EPA, 1977) at least every three years to provide a forum for the upgrading and improvement of modeling in support of the Act. Since it published the modeling guideline in April 1978 (EPA, 1978), EPA has conducted studies to identify various aspects of the performance of air quality simulation models. However, model performance standards still do not exist.

The criteria used to select models for recommendation in the 1978 Guideline were basically the availability of documentation and staff experience with the algorithms and software. Where a number of possible models could be used for the same or similar applications, factors such as model availability and familiarity to the user community were considered. Because there was no consensus about model performance standards, model availability and familiarity had a greater influence on model selection.

A logical basis for selection of models should emphasize how well a model compares to a wide range of data it purports to simulate. It should be tested on how well it simulates generally accepted principles of atmospheric behavior. This paper recommends measures to be used to evaluate the air quality performance of simulation models and outlines some approaches for developing model performance standards.

3. Measuring model performance

Workshop participants were asked to review a number of reports and publications dealing with performance evaluation (MacKay, 1980; Brier, 1975; Bowne, 1980; Hayes, 1979; Hillyer *et al.*, 1979; Venkatram, 1979; Bornstein, 1980). It was recommended that the distinction between "scientific" and "operational" performance measures be considered. "Scientific" implies at least some understanding of cause-and-effect relationships that rely on testing model components and extensively detailed data collection. "Operational" implies a comparison with data exclusively within a particular application context. Examples of the former are the EPRI Plume Model Validation study (EPRI, 1978) and the EPA Complex Terrain study (Holzworth, 1980). An example of the latter is the comparison of peak observed concentrations with peak measured concentrations, without regard to the meteorological conditions, which is often employed in the regulatory setting. The group was charged to consider that 1) meteorological events on an hourly basis often are not independent of one another; 2) air quality standards are written for 1, 3, and 24 h and annual averaging periods, so performance should be judged accordingly; 3) conservatism is often built into models; 4) meteorological and emissions inputs contain inherent uncertainties with resulting implications about the limits of model performance; and 5) model performance may be different for different applications.

It is important to distinguish the concepts of verification, evaluation, and validation of air quality models. Verification is proof of the accuracy, reality, or truth of a model. Evaluation is the process of examining and appraising the performance by comparing the model's concentration estimates to measured air quality data. Validation is the establishment of a conclusion by detailed and copious evidence that leads to formal recognition. This may include several evaluations. Models are evaluated in the hope that they will be validated.

Another distinction to be made is between point source models and urban models. Urban model evaluations generally include a large amount of data from a wide distribution of sources. In contrast to relatively uniform concentrations over the urban area, the pollutant concentrations associated with single point sources tend to exhibit very sharp variations. A complicating factor in the point source case is that the plume centerline (maximum concentration) is measured only intermittently by monitors associated with the source; consequently the data set tends to contain a majority of zero values. Another issue is the inherent difference between the concentration fields of primary and secondary pollutants.

a. General considerations

1) USING MODELS FOR REGULATORY DECISIONS

It was emphasized that models are both necessary and desirable tools to help make equitable decisions in air quality management. It should be recognized, however, that a model produces a result that represents a predicted value of a response at a particular time and place. The actual value of the response produced by the natural system generally will differ from the predicted value; therefore, it is important for a modeler to provide the decision-maker with not only a predicted value, but also an estimate of the error associated with that value. Both the predicted value and its estimated error can then be combined to provide a confidence interval for the true responses.

2) STATISTICAL ASPECTS OF AIR QUALITY GOALS
The nature of current National Ambient Air Quality Standards and Prevention of Significant Deterioration increments requires that an extreme concentration value (highest, second highest) be estimated by models. Workshop participants agreed that the accuracy of highest or second highest estimates is expected to be poor and difficult to evaluate statistically. Statistical evaluations have greater meaning when applied to a relatively larger number of values than to one or two extremes. Generally, statistical evaluations applied to an upper percentile (2% or 5%) of the predicted values are more informative than those applied to only the highest or second highest prediction. Evaluations applied to estimates of mean performance will supply more information about overall model performance than will evaluations applied to extremes only. The workshop participants, therefore, recommended that the statistical form of standards and increments be changed to consider the upper 2%–5% of concentrations rather than one or two extreme values.

3) SCIENTIFIC JUDGMENT AND EVALUATION
Statistical measures should not be used exclusively to evaluate air quality simulation models. A scientific evaluation, involving examination of cause-and-effect relationships and extensively detailed data collection, should be included.

Many participants felt strongly that scientific judgment might prove to be the only effective mechanism to distinguish between models. The complexity of the physical system requires close examination of how accurately a model represents real conditions in that system. Because concentration is based on atmospheric turbulence, the dispersion process is unlikely to be understood in detail, forcing reliance on ensemble averages. Model predictions based on such averages differ from measurements. Often, sufficiently detailed and lengthy data are not available to conduct a meaningful evaluation. In such cases, a value judgment with regard to the physical representation of specific components of a model may prove the only critical mechanism.

A model should be examined to determine whether its underlying assumptions are reasonable from a scientific standpoint. The model assumptions should be consistent with physical principles, such as conservation of mass and the second law of thermodynamics, and generally accepted atmospheric behavior. The algorithms and parameters adopted in the model to describe dispersion and other processes must be suited to the application (Hanna *et al.*, 1977). The computational technique used in the model should not contribute to errors in its predictions. The model computer code should be consistent with the assumptions set forth in the model user's manual or other technical support document. Quality control of model codes should be emphasized. A final component of scientific evaluation is an analysis of the sensitivity of the model to variation in input parameters.

4) DATA USE
Data used for evaluating air quality models should be representative of the conditions the model is designed to simulate. The data should be accurate, precise, and complete. They should be gathered at places and over time intervals to which the model pertains. For many statistical procedures, it is desirable to maintain independent data sets for development and for evaluation. Model evaluation should be conducted using data not used in developing the model. For example, any background concentration value that may be established, associated with a point source monitoring network, in a location with measurable ambient concentration levels should be retained in the data. Data should be used in their dimensional form.

5) DATA INDEPENDENCE
Air quality data at any specific location commonly exhibit correlation associated with the persistence of meteorological events. This fact led to discussion concerning the importance of dealing with dependent (time-correlated) versus independent (not time-correlated) data sets. In many standard statistical tests, an assumption of independence is made. For this reason, some preprocessing of the time series of data points may be necessary in order to use properly statistical techniques that assume independent data points. The preprocessing techniques would attempt to eliminate characteristics of dependent measurements from data streams. This is most important when using tests of the statistical significance of model performance measures. If a strong temporal correlation exists in both the observations and the model predictions, then the statistical significance associated with the measure of model performance may be overestimated unless the correlation is removed.

For example, typical hourly averaged concentration data often exhibit a strong diurnal pattern. A good model prediction should duplicate this pattern. To evaluate model performance, the differences between the observations and the prediction also should be considered. These differences should not show a diurnal pattern. If they do, it suggests inadequacies in the model. In any case, if the differences are not independent, statistical tests should be applied to a reduced set of the differences that are independent (perhaps every fifth hour or every twenty-second hour, etc.).

6) DATA NORMALIZATION
More informative statistical measures can be applied to data that are normally distributed than to data that are not. For this reason, it is generally recommended that data used for model evaluation be transformed so that their random variations approximate a normal distribution. Various distributions (exponential, lognormal, etc.) are exhibited by point source and urban air quality data under special conditions (Larsen, 1971; Gifford, 1972). Therefore, it may be advantageous for evaluation purposes to transform data to develop a distribution that is more nearly normal than the original distribution.

7) AVERAGING TIME
Because standards are written for various averaging times (3 and 24 h and annual averages) it will be necessary to evaluate model performance as a function of

these averaging times. The 1 h averaging time, however, is the basic element for evaluation.

8) PREDICTABILITY AND ACCURACY

Uncontrollable stochastic variations of air quality and meteorological elements impart reduced predictability to air quality simulation models. Workshop participants recognized four other areas of limitation in the accuracy of air quality model estimates:

1) the quality of the emissions data;
2) the quality (appropriateness or relevance to the problem) of the meteorological data;
3) the quality (appropriateness or relevance to the problem) of ambient air quality measurements available for comparison to predictions;
4) the capability of the algorithms (model) to reproduce the natural events.

These considerations suggest a practical limit to the accuracy that a model may attain (Venkatram, 1979). The limit of predictability varies for different applications, notably for different averaging times and topographical settings.

9) ERROR AND DIFFERENCE

The term difference is used in this summary to describe differences between observed and calculated values; they may result from various types of errors. The errors may be sampling errors caused by incompleteness of the sample coverage, instrumental errors resulting from imperfections in the measurement technology, computational errors caused by the inadequacy of the numerical approximation of an analytical expression, they may result from the use of an inappropriate model, or they simply can be associated with the natural, random nature of the problem. Ideally, the sampling and instrumental errors in the input data can be minimized so that the model errors (physical parameterization and computational approximation) can be evaluated.

b. Measures of performance

Considerable effort has gone into identifying measures of air quality model performance (Hayes, 1979; Londegren, 1980; Javitz and Ruff, 1979; Mills *et al.*, 1980; Bencala and Seinfeld, 1979; Nappo, 1974). Three groups of both predicted and observed data for which performance measures should be developed were identified by the workshop participants:

1) observed and predicted concentration field values paired for a particular location in space at a particular time;
2) a peak concentration data set with various degrees of time and space pairing;
3) cumulative frequency distributions of the unpaired (in time) observed and predicted concentration values.

Analysis of the paired data sets was considered to represent the most stringent test of a model, because the model is required to simulate all concentrations at a particular place and time. Analysis of the peak concentration data provides a somewhat less stringent test, because the model is asked only to perform under worst-case conditions, which represent a subset of all conditions.

1) SPECIFIC PERFORMANCE MEASURES

Two general types of performance measures were identified: 1) measures of difference (often termed residuals), and 2) measures of correlation. Difference measures represent a quantitative estimate of the size of the differences between observed and predicted values. Correlation is the quantitative measure of the association or agreement between predictions and observations. Because the statistical theory is more highly devel-

TABLE 1. Measure of performance. Numbers in parentheses refer to corresponding equation numbers in Appendix—Notes on Table 1.

	Difference				Correlation		
			Gross Variability (or Gross error)				
Concentration	Average (or bias)	Variance (or noise)	Standardized	Aver. absolute	Temporal	Spatial	Space/Time
Paired values C_0, C_p	$\bar{d}$ (2)	S_d^2 (3)	MSE_d (4)	$\lvert\bar{d}\rvert$ (6)	$r(\Delta T)$ (11)	r_s (11)	r (11)
Peak values maximum observed $C_0(L_{(n)}, T_{(n)})$:							
$-C_p(L_{(n)}, T_{(n)})$	$\bar{d}_M(L_{(n)}, T_{(n)})$ (7,2)	$S_{\bar{d}_M}^2$ (7,3)	$\mathrm{MSE}_{\bar{d}_M}$ (7,4)	$\overline{\lvert d_M(L_{(n)}, T_{(n)})\rvert}$ (7,6)	$r_{L_{(n)}}(\Delta T)$ (7,11)		
$-C_p(L_j, T_{(n)})$	$\bar{d}_M(L, T_{(n)})$ (8,2)	$S_{\bar{d}_M}^2$ (8,3)	$\mathrm{MSE}_{\bar{d}_M}$ (8,4)	$\overline{\lvert d_M(L, T_{(n)})\rvert}$ (8,6)	$r_L(\Delta T)$ (8,11)	$\dot{\Delta}_{(n)}(L, T_{(n)})$ (12)	
$-C_p(L_{(n)}, T_j)$	$\bar{d}_M(L_{(n)}, T)$ (9,2)	$S_{\bar{d}_M}^2$ (9,3)	$\mathrm{MSE}_{\bar{d}_M}$ (9,4)	$\overline{\lvert d_M(L_{(n)}, T)\rvert}$ (9,6)	$\Delta T_{(n)}$ (14)		
$-C_p(L_j, T_i)$	$\bar{d}_M(L, T)$ (10,2)	$S_{\bar{d}_M}^2$ (10,3)	$\mathrm{MSE}_{\bar{d}_M}$ (10,4)	$\overline{\lvert d_M(L, T)\rvert}$ (10,6)	$\Delta T_{(n)}$ (14)	$\dot{\Delta}_{(n)}(L, T)$ (13)	
Frequency distributions $C_0(T)$ and $C_p(T)$	t, z Wilcoxon/ Mann-Whitney	F, χ^2	χ^2, Kolomogorov-Smirnov				

oped, difference measures seem to be better suited to quantitative performance evaluation.

The following measures of difference were suggested:

1) the bias (average) of the difference (observed −predicted);
2) the variance of the difference (noise);
3) the gross variability (gross error) of the difference.

These are actually related measures, because the sum of the square of the bias plus the variance equals the square of the gross variability. The following measures of correlation were suggested:

1) time;
2) space;
3) time and space combined.

The first is a measure of the association between observed and predicted values over time, the second is a measure of the spatial association between those values, and the third is a combination of the first two. Although correlation coefficients can be based on assumed linear or nonlinear relationships, for simplicity our consideration is limited to linear ones.

The measures of difference and correlation can be applied to the three groupings of data identified previously to produce a table of performance measures (Table 1). The table presents the three statistical functions (bias, noise, and gross error) each applied to the full field paired (in time) concentration data, the selectively paired peak concentrations, and the full field unpaired (in time) concentrations.

2) GRAPHIC DISPLAYS

Informative graphic techniques should be included in any performance evaluation. Histograms and displays of cumulative frequency distributions are appropriate for all the difference measures. Scatter diagrams, isopleths, and time series displays are useful for correlation measures.

c. Using performance measures

The statistically based performance measures listed in Table 1 are rather comprehensive. Not all measures are meaningful for all problems. This section considers some issues relative to the qualification of performance evaluation, especially as applied to point source models.

1) PRELIMINARY ANALYSES

One step in evaluation is to examine whether the model is able to account for a selected number of the highest (peak) concentration measurements. Because this is particularly difficult for point sources, it is recommended that a case study approach be adopted.

A case study approach should include consideration of the meteorological conditions associated with the highest 10–25 concentration occurrences. For a particular monitoring location, the comparison of meteorological characteristics, such as stability and wind speed class, is recommended. It should consider averaging times of 1 h as well as averaging times of importance to regulatory standards (e.g., 3 h and 24 h periods for SO_2).

The case study approach can identify problems with a model that might not be apparent using statistical performance measures nonselectively. For example, assuming a Gaussian plume model, if high measured concentrations were found during a period for which the model predicted zero values everywhere, then it might be suspected that the model's treatment of mixing height penetration or the value of the plume, σ_z, is in error. Similarly, if model results indicated that most of the highest concentrations would occur with slightly unstable conditions, while the highest concentrations actually occurred with very unstable conditions, then probably either the method used for the assignment of the stability or the choice of dispersion curves associated with these determinations is in error. More generally, if a case study approach identifies problems with the model assumptions, then the model can be changed. The degree of interpretation and conclusion derived from such analysis depends upon the confidence placed on the accuracy and representativeness of the model input data and the measured concentrations.

2) DATA BASE CONSIDERATIONS

The value of quantitative performance measures depends on data quality, data base size, and the number of monitors in a network. For a small network (approximately five monitors or fewer is the most common regulatory situation), models generally are evaluated only for their ability to predict the concentrations. For a large network (greater than 10 or of order 100 or more monitors for field studies), models can be evaluated for both predictions of concentrations and their spatial alignment. The distinction between small and large networks is somewhat subjective and must be judged on a case-by-case basis.

a) Small network comparisons by meteorological category.—For this case, a useful performance attribute is the difference between observed and predicted concentrations as a function of different meteorological categories. There should be enough categories to encompass a wide range of meteorological conditions, and the amount of data in each category should be large enough for a statistically meaningful comparison of observed and predicted values. For example, the data could be stratified by lumped stability class groups (A–C, D, E–F) instead of by the individual classes (A, B, C, D, E, and F), in order to develop a sufficiently large number of observations. Individual observations and predictions should be compared for a particular meteorological category only if the wind vector is within a specified angular sector about the line joining the source and monitor. Major uncertainty often exists regarding crosswind location of the plume centerline and the monitor; therefore, prediction methods must simulate the effects of wind direction uncertainty. This can be done by using random numbers to compute simulated wind directions within the angular sector. A second method involves calculating "sector-averaged" values to compare to the mean of the observed values, and calculating the corresponding expected standard deviation of these values to compare to the standard deviation of the observed values. Statistical tests (Student's t or Wilcoxon) could be used to indicate whether the average observed concentration and the average predicted concentration for a particular meteorological category were significantly different from one another. An F-test could be used to compare the estimated variances of the observed and predicted concentrations for each category.

The peak predicted and peak observed concentrations in each meteorological category also can be compared when the wind is in the appropriate sector. Predicted concentrations could be computed at the downwind distance of the monitor for an array of across-plume locations. The highest value of this spatial array of calculated concentrations then could be selected to compare to the 10–25 highest observed concentrations. A Wilcoxon or other appropriate test could then be used to determine statistical significance.

b) Small network comparisons regardless of meteorological category.—In some cases, the quality and/or representativeness of meteorological data and the number of available monitoring sites may not allow a stratification of concentrations by meteorological category. Workshop participants agreed that it is difficult, if not impossible, to evaluate a model under these circumstances. A major factor in any such evaluation would be the concurrence of measured and modeled meteorological conditions that lead to maximum concentration predictions. If maximum concentrations are predicted to occur under meteorological conditions different from conditions found with actual maximum concentrations, the model is not believable.

c) Small network averaging time considerations.—To evaluate model predictions for 3 h averaging time, meteorological data based on the hour showing the greatest impact in the 3 h period should be used for the calculation. In the case of a 24 h model, days should be classified as to wind persistence, season, and most frequent daytime stability. A case study analysis is recommended for the evaluation of peak 3 h and 24 h average concentrations. For 3 h and 24 h model evaluations, an additional performance measure is the frequency distribution of peak to average (peak 1–3 h average and peak 1–24 h average) concentration ratios. If the model predicts the 1 h values well, it is an indication of the ability of the model to simulate multiple-hour concentration averages, especially if the ratios are stratified by critical meteorological conditions. Caution is advised because the statistical properties of a peak to mean ratio may be difficult to analyze.

d) Small network limitations.—Although these data base considerations are an attempt to improve model evaluations by using the existing data resources, they have limitations. Models tested on such data sets should not be used for circumstances very different from the specific study.

e) Large network.—The measures listed in Table 1 are most useful when applied to the relatively well described concentration field sampled by a large network. Generally, such data sets consist of tracer releases rather than actual pollutants. In field studies, the fixed network is occasionally supplemented by mobile and remote monitoring capabilities. This large data situation is preferred. Efforts to evaluate models should be based upon such data sets.

4. Setting performance standards

A rational approach to model selection based upon performance evaluations, preferably in a precise statistical setting, is necessary. However, there is a lack of experience with performance measures, an inadequacy of data bases to provide critical test material, and the potential for any standard to be misused. Although regulatory applications of models define some requirements on performance, how accurate models need to be remains a major policy issue. How accurate the models are capable of being is a different issue. Modelers should convey to model users an indication of the validity, accuracy, or, less precisely, believability of their model's predictions.

It is not appropriate to establish specific numerical standards for model performance. Insufficient experience with rigorous model evaluation, various potential uses, and scientific limitations on concentration predictability suggest caution in judging model performance.

The scientific limit of predictability represents the best a model can do. Practically, models probably cannot achieve this. Neither is it possible to know quantitatively how closely a model may approach optimum. Nevertheless, identification of a model's scientific limits for a specific application could establish whether or not such a model is reasonably applied to that problem. This information would help in selecting general model types for specific applications.

a. Comparison with EPA-approved models

As an alternative to setting standards, those workshop participants considering point source models proposed that the measures of performance specified in Table 1 be used to compare models. An evaluation procedure should include several steps:

1) Run models described in the *Guideline on Air Quality Models* (EPA, 1978) with the best possible data base.
2) Apply the measures of performance listed in Table 1 to provide reference levels of performance. It is important to emphasize that the reference levels of performance are not a standard of performance. They are a movable but defined reference level analogous to zero degrees Fahrenheit. The reference level is arbitrary, unlike zero kelvin, which is absolute.

3) Compare the outputs of new or candidate models with outputs of the appropriate EPA (reference) models, using the performance measures in Table 1.
4) Test for the statistical significance of the differences between models.

Following these procedures would have several advantages: a profile of the model's performance will develop, it will be found that some models will perform more satisfactorily for some applications than for others, and site-specific and time-preferable models should become apparent.

Models may be identified that perform better over a broader spectrum of applications and for a broader spectrum of measures than the appropriate guideline model. They should be adopted as a replacement to the guideline model. A continuous procedure to upgrade models can evolve.

The process and procedure of evaluating models against each other will provide incentives for the scientific community, not only to develop better and more applicable models, but also to provide the air quality manager with more specific information on the levels of confidence associated with each model's data.

Since guideline models are used in formal regulatory activities, they are in a sense operationally acceptable. The acceptability of alternative models can therefore be determined through model performance profile comparisons.

b. Use of statistical confidence intervals

The participants considering urban models took the following initial positions regarding appraisal of model performance: 1) no adequate basis currently exists for setting performance standards; nevertheless, 2) decision-makers require a rational analytical framework in which to interpret the statistics derived from performance measures. It was finally concluded that 1) the use of statistical acceptance/rejection tests for the performance measure appraisal seems inappropriate in light of large uncertainties associated with model inputs and formulation, and limitations imposed by most observational data sets; therefore, 2) it is reasonable to urge that the use of statistical confidence interval statements be accepted as a common tool in model performance evaluation and as a recommended initial step towards setting performance standards.

A confidence interval statement for a parameter is more informative than a statistical test. The interval statement provides information relative to many different hypotheses, whereas the statistical acceptance/rejection test is a decision rule about a single hypothesis. The acceptance/rejection information for any single hypothesis is contained in the interval statements. Judgments concerning model performance can be made using confidence interval statements along with other relevant information.

Interval statements can be constructed for the bias and for the variance using the Student's t and the chi-square distribution. For the estimated bias of the difference, $\bar{d}$, such a statement can be approximated by

$$\bar{d} \pm t_v \left(\frac{S_0^2}{N} + \frac{S_p^2}{N} \right)^{1/2}$$

where t_v is the Student's t statistic for a data trace with v degrees of freedom (related to the number of independent observations, N) and a selected value of the confidence interval, S_0^2 is the estimated variance of the observations, and S_p^2 is the estimated variance of the predictions.

If S_d^2 is the estimated variance of the differences, the confidence interval for the true variance, σ_d^2, can be expressed as a probability

$$\frac{vS_d^2}{\chi_v^2, 1-x} < \sigma_d^2 < \frac{vS_d^2}{\chi_v^2, x}$$

where χ_v^2, x is the x^{th} percentile of the chi-square distribution for v degrees of freedom.

Interval statements alone cannot be used to determine acceptability of models. In fact, they will suggest rejection only when model performance is quite bad. Accumulation of sufficient data and many analyses are needed to provide sufficient evidence for acceptance or rejection of models.

If the assumptions concerning the use of the distribution upon which the interval construction is based are seriously violated, the interval statement itself will be inaccurate. However, the confidence interval represents the strongest statement that the data and statistics procedures permit. It should be pointed out that any confidence interval statement presumes the selection of a level of the confidence (probability).

Because of the nature of the atmospheric system, establishment of confidence limits and testing for the statistical significance of differences between models will very likely be of only limited initial success. The statistical apparatus recommended is designed to establish whether or not the observations and the model predictions come from the same population of data. We know they do not. In fact, the infinite population of a typical meteorological variable, when sampled from one occasion to the next or from one location to the next, generally does not show much evidence of being from the same population. This suggests that the statistical significance of model performance measures will not be established easily. The question facing the regulatory decision-maker is whether the model predictions reasonably approximate actual situations. The answer to this will come only with experience in the application of the suggested performance measures on continuously improving data bases. Thus, there is a critical need for good data bases and for additional research.

5. Research needs

Because there is so little experience with these techniques, there is a great need for research to generate new data, analyze data, and develop and test new performance measures. Requirements range from fundamental academic research in which statistical techniques and diffusion models are developed to applied research in which field monitoring studies are conducted.

Potential benefits from improved accuracy in the prediction of concentrations resulting from pollution emissions, and hence from improved decision-making, are great. Fundamental problems of atmospheric physics and chemistry currently limit the ability to predict the consequences of pollutant release. As performance measures are refined, one anticipated outcome is the realization that the models, while significant, are primitive. Considerably more fundamental attention should be devoted to their basic physics and chemistry.

a. Development and description of performance measures

Further work must be done to determine needs for and applications of performance measures. The relevance of each measure to specific models and applications should be established. A handbook of performance measures should be developed, which should contain statistical techniques to be used to evaluate model performance and standardize symbols and definitions. A computerized statistical package could be included with the handbook.

Independent research should be done to develop new ways of applying standard statistical techniques to this problem. Experience from comparing actual data in operational models will add knowledge of the information conveyed by the different measures. With this experience, an understanding of appropriate weighting schemes for evaluating performance in various applications will evolve. The temporal characteristics of both the observed and predicted data traces are of particular concern. Research is needed to adapt statistical time series analysis methods to model validation.

b. Application of performance measures to point source models

Recommendations are needed on the kinds of experimental programs providing the best plume concentration measurements for testing models. High priority should be given to those research programs that overcome uncertainties in monitor location relative to the plume centerline. Two of these programs are a very dense network of monitors, such as used in the EPRI program, and the mobile van technique, where an instrumented van makes repeated transects across the plume centerline to define average crosswind concentration profiles (Weil, 1979).

The quantitative variation of performance measures with the number of monitoring stations used in the network should be tested. For example, with the EPRI data set performance measures can be computed using the entire 200 samplers in the network and then compared to performance measures using only a small subset of 5–10 monitors. This will help evaluate how the size of the data set affects the ability to evaluate a model.

A technique is needed to estimate the upper percentiles of the concentration frequency distribution for a particular meteorological condition or category. Models are capable at best of predicting the ensemble average concentration for a particular meteorological condition, but provide little information on the variation about the predicted average. This variation could be obtained by using an empirically determined distribution (possibly log normal) and an empirically determined standard deviation from that distribution. The mathematical function descriptive of the distribution should be the same for different meteorological conditions, but, for example, the standard deviation could be quite different for various stabilities. Such a natural variability could be taken into account when estimating the upper percentiles of the concentration frequency distribution at a particular receptor over a long period of time.

c. Analysis of characteristics of meteorological data

The accuracy of an air quality model depends partly on the variability and representativeness of meteorological data. This information can be used to assess the sensitivity of models to errors in input data. Existing meteorological network data (e.g., Regional Air Pollution Study (RAPS), Idaho Falls, etc.) should be used to determine variability resulting from natural effects and instrument capabilities. Because it is especially important that models accurately reflect different areas and time periods, this representativeness should be investigated in different types of terrain. It will be necessary to create new data sets to cover all the conditions examined. Also, the number of instruments needed and location of the sites to provide representative values of wind velocity or stability need to be determined.

d. Analysis of characteristics of air pollution monitoring data

Several air pollution monitoring studies (e.g., Sulfate Regional Experiment (SURE), Large Power Plant Effluent Study (LAPPES), Regional Air Pollution Study (RAPS) have not been fully analyzed. These data sets should be used to determine the natural variability (for example, the bias and the estimated standard deviation) of concentrations under narrow ranges or categories of meteorological conditions and source emissions. For example, one category might be D stability, wind sector $\Delta\theta$ about the source-monitor vector $10^\circ < \Delta\theta < 30^\circ$, and wind speed $4\ \mathrm{m/s} < u < 6\ \mathrm{m/s}$. To have enough data in each category, thousands of data points are needed. The characteristics of the distribution function (log normal, normal, exponential, Gamma, etc.) should be determined, especially for the upper percentiles (highest concentrations). Confidence intervals should be constructed for the characteristics of the distribution functions. Research with these same data sets could be conducted to determine how to produce normally distributed, independent data sets. Sample size requirements must be set. Trends and obvious harmonic functions (e.g., diurnal curves) should be removed using time series analysis. The autocorrelation function can be estimated. The resulting independent, normal data set would provide an example for application of powerful, statistical measures.

These data sets should be used to study peak concentration episodes, especially for isolated point sources. They are useful in determining both spatial patterns (use quantitative measures such as in Table 1) and the meteorological conditions that cause extreme concentration peaks. Long sampling times (>1 year) should be used to increase chances of finding combinations of meteorological conditions that lead to extreme peaks.

New data sets should be obtained in different types of terrain and for different combinations of sources (tall point buoyant, short stack nonbuoyant, multiple point source, urban area source). Various averaging times should be studied using these analysis techniques.

e. Evaluation of diffusion models

Several models should be run using good data in order to produce predicted concentration estimates to compare with observations. The statistical tests described previously and

Scenes of Woods Hole, Mass., site of the AMS Workshop on Dispersion Model Performance, 8–11 September 1980. (Photos courtesy of Robert Bornstein.)

listed in Table 1 then should be applied to the observed and predicted data sets. Typical expected values and confidence intervals for these statistical measures should be estimated for different types of models (box, grid, Gaussian) and for different categories of meteorological conditions. Space and time patterns of statistical measures for peak concentration episodes should be studied. Results for new models should be compared to results for representative models from the guidelines document (EPA, 1978).

The scientific assumptions in models should be analyzed to find out why they produce certain behavior in the statistical performance measures. For example, a poor default assumption for calm winds may lead to poor estimates of peak episodes. It may be necessary to test modules of the model, such as the plume rise formula or the dispersion parameterization, separately. Special data sets may be needed for these tests.

Model sensitivity to variations in the inputs emissions data or meteorological conditions should be studied using the typical variations calculated from the studies mentioned previously. The effect of these natural variabilities on confidence intervals for statistical measures (such as the cumulative distribution) should be estimated.

6. Conclusions

A set of statistics has been proposed that can provide a rational framework for quantitatively evaluating the nature of differences between observations and predictions by models. Statistics was suggested as a tool to provide confidence in model predictions, as well as to compare new models against those models recommended by EPA in its *Guideline on Air Quality Models.* However, because of extremely limited experience with these measures of performance, all the models should be tested using the framework suggested in this paper.

Appendix—Notes on Table 1

C is not necessarily the concentration value itself but could be a transformed value. This transformation would be applied to attempt to obtain a normal distribution of values.

d is the difference in the concentrations, evaluated from a pair of concurrent observed and predicted concentrations via the following definition, d:

$$d = C_0(x,T) - C_p(x,T) \tag{1}$$

where C_0 and C_p are the original or transformed observed and predicted values, respectively.

$\bar{d}$ estimates the bias of the concentration difference defined as:

$$\bar{d} = \frac{1}{N} \Sigma\, d \tag{2}$$

where N is the number of observations. The bias is the first moment of the distribution of differences.

S_d^2 is the estimated variance of the concentration difference and is calculated from the following definition:

$$S_d^2 = \frac{1}{N-1} \Sigma\, (d - \bar{d})^2 \tag{3}$$

The variance is the second moment of the distribution of differences. The square root of the variance is the standard deviation.

The mean square error[2] (MSE) of the concentration difference is evaluated from the following definition:

$$\mathrm{MSE}_d = \frac{1}{N} \Sigma\, d^2 \tag{4}$$

The root mean square error, rms, is the square root of MSE.

The mean square error is related to the bias, $\bar{d}$, and the variance, S_d^2, by:

$$\mathrm{MSE} = \frac{N-1}{N} (S_d^2) + \bar{d}^2 \tag{5}$$

For normally distributed variables, the bias has a normal distribution, while the variance S_d^2 has a chi-squared distribution. The mean square error has a compound distribution.

$|\bar{d}|$ is defined to be the average absolute gross error of the concentration difference evaluated from the following definition:

---

[2] The general definition of MSE involves a reference level about which MSE is calculated so the $((N-1)/N)S_d^2$ is MSE about the average. The MSE defined here is about a reference level of zero.

$$|\bar{d}| = \frac{1}{N} \Sigma |d| \tag{6}$$

This statistic is more robust (less affected by removal of a few outliers) than is rms.

$d_M(L_{(n)}, T_{(n)})$ is the difference between the observed maximum concentration at a given time over the entire spatial field and the corresponding predicted value at the same location and time that the maximum occurred, as defined by:

$$d_M(L_{(n)}, T_{(n)}) = C_0(L_{(n)}, T_{(n)}) - C_p(L_{(n)}, T_{(n)}) \tag{7}$$

where the subscript M denotes a maximum concentration and $L_{(n)}$ is the location of the site with the nth maximum observed concentration at time $T_{(n)}$, i.e., $n = 1$ is the highest, $n = 2$ the second highest, etc. This is the most stringent of the three recommended measures of the difference d_M.

$d_M(L, T_{(n)})$ is the difference between the observed maximum concentration at a given time over the entire spatial field and the maximum predicted value in the entire spatial domain at the same time, as defined by:

$$d_M(L, T_{(n)}) = C_0(L_{(n)}, T_{(n)}) - C_p(L_i, T_{(n)}) \tag{8}$$

where L_i is the site of the maximum predicted concentration at time $T_{(n)}$. Note that in this case the maximum observed and predicted values do not have to occur at the same location, but are at the same time.

$d_M(L_{(n)}, T)$ is the difference between the observed maximum concentration at a given time over the entire spatial field and the predicted maximum concentration at that site anytime during a time period, as defined by:

$$d_M(L_{(n)}, T) = C_0(L_{(n)}, T_{(n)}) - C_p(L_{(n)}, T_i) \tag{9}$$

where T_i is the time of the predicted maximum at $L_{(n)}$. Note that in this case, the maxima occur at the same site, but not necessarily at the same time.

$d_M(L, T)$ is the difference between the observed maximum concentration at a given time over the entire spatial field and the predicted maximum concentration at any time and location, as defined by:

$$d_M(L, T) = C_0(L_{(n)}, T_{(n)}) - C_p(L_j, T_i) \tag{10}$$

where $C_p(L_j, T_j)$ and $C_0(L_{(n)}, T_{(n)})$ are the maximum predicted and observed concentrations, respectively. These can occur anywhere in the region at any time over the appropriate time interval. This is the least stringent of the four recommended maximum difference measures.

$\overline{d_M(L_{(n)}, T_{(n)})}$, $\overline{d_M(L, T_{(n)})}$, $\overline{d_M(L_{(n)}, T)}$, and $\overline{d_M(L, T)}$ are the estimated biases of the maximum concentration differences, respectively, as computed from Eq. (2).

$S^2[d_M(L_{(n)}, T_{(n)})]$, $S^2[d_M(L, T_{(n)})]$, $S^2[d_M(L_{(n)}, T)]$, and $S^2[d_M(L, T)]$ are the variances of the maximum concentrations estimated, respectively, as computed from Eq. (3).

MSE$(L_{(n)}, T_{(n)})$, MSE$(L, T_{(n)})$, MSE$(L_{(n)}, T)$ and MSE(L, T) are the mean square errors of the maximum concentration differences, respectively, as computed from Eq. (4).

$\overline{|d_M(L_{(n)}, T_{(n)})|}$, $\overline{|d_M(L, T_{(n)})|}$, $\overline{|d_M(L_{(n)}, T)|}$, and $\overline{|d_M(L, T)|}$ are the average absolute gross errors of the maximum concentration differences, respectively, as computed from Eq. (6).

$r(\Delta T) = f[C_0(L_{(n)}, T), C_p(L_{(n)}, T_{(n)})]$ is the cross-correlation coefficient. It is calculated using observed and predicted paired values at one site, $L_{(n)}$, at all positive and negative time lags (ΔT) from the following definition of the correlation coefficient, r:

$$r = \frac{\Sigma (C_0 - \overline{C_0})(C_p - \overline{C_p})}{\left\{\Sigma (C_0 - \overline{C_0})^2 \cdot \Sigma (C_p - \overline{C_p})^2\right\}^{1/2}} \tag{11}$$

$r_s = f[C_0(L, T_{(n)}), C_p(L, T_{(n)})]$ is the spatial correlation coefficient. It is calculated using observed and predicted paired values in Eq. (11) at zero time lag and at a given time ($T_{(n)}$) using data for all of the sites together.

$r_{L(n)}(\Delta T)$ is a cross-correlation coefficient calculated from Eq. (11) using observed and predicted maximum concentrations used in calculating $d_M(L_{(n)}, T_{(n)})$ at all positive and negative time lags (ΔT). This statistic defines the time lag that yields the best correlations between observed and predicted maximum concentrations at a single location.

$r_L(\Delta T)$ is a cross-correlation coefficient calculated from Eq. (11) using observed and predicted maximum concentrations used in evaluating $d_M(L, T_{(n)})$ at all positive and negative time lags (ΔT). This statistic defines the time lag that yields the best correlation between observed and predicted maximum concentrations without regard to the position of the maximum in space.

$r(0)$, $r_{L(n)}(0)$, and $r_L(0)$ are the cross-correlation coefficients at zero time lags. These are the versions of $r(\Delta T)$, $r_{L(n)}(\Delta T)$, and $r_L(\Delta T)$ usually computed in evaluation studies.

$\vec{\Delta}_{(n)}(L, T_{(n)})$ is the vector difference between the predicted and observed maximum (highest $n = 1$, second highest $n = 2$, etc.) concentrations at time $T_{(n)}$ anywhere in the spatial field. It is used with $d_M(L, T_{(n)})$, and is evaluated from:

$$\vec{\Delta}_{(n)}(L, T_{(n)}) = \vec{D}_0(L_{(n)}, T_{(n)}) - \vec{D}_p(L_i, T_{(n)}) \tag{12}$$

where $\vec{D}_p(L_i, T_{(n)})$ and $\vec{D}_0(L_{(n)}, T_{(n)})$ are the vector distances in two components from some origin to the location of the predicted and observed maximum concentrations at time $T_{(n)}$, respectively.

$\vec{\Delta}_{(n)}(L, T)$ is the vector difference in two components between the predicted and observed maximum concentrations anywhere in the temporal and spatial fields. It is used with $d_M(L, T)$ and is evaluated from:

$$\vec{\Delta}_{(n)}(L, T) = \vec{D}_0(L_{(n)}, T_{(n)}) - \vec{D}_p(L_i, T_j) \tag{13}$$

where $\vec{D}_0$ and $\vec{D}_p(L_i, T_j)$ are the vector distances in two components from some origin to the location of the maximum observed concentrations and the maximum predicted concentrations anywhere in the region at any time over the appropriate time interval.

$\Delta T_{(n)}$ is the time difference from the maximum (highest $n = 1$, second highest $n = 2$, etc.) observed concentrations using data from the entire fields. It is used with both $d_M(L_{(n)}, T)$ and $d_M(L, T)$ and is computed from:

$$\Delta T_{(n)} = T_{0(n)} - T_{p(n)} \tag{14}$$

where $T_{p(n)}$ and $T_{0(n)}$ are the times of the predicted and observed maximum concentrations, respectively. These can occur anywhere in the region at any time over the appropriate time interval.

References

Bencala, K. E., and J. H. Seinfeld, 1979: An air quality model performance assessment package. *Atmos. Environ.*, **13**, 1181–1185.

Bornstein, R. D., 1980: Survey of statistical techniques used in validating air quality simulation models, 30 pp. (Unpublished manuscript.)

Bowne, N. E., 1980: Validation and performance criteria for air quality models. *Conference Papers, Second Joint Conference on Applications of Air Pollution Meteorology* (*New Orleans*), AMS, Boston, pp. 614–626.

Brier, G. W., 1975: Statistical questions relating to the validation of air quality simulation models. *EPA-650/4-75-010*, USEPA, Research Triangle Park, N.C., 22 pp.

Electric Power Research Institute, 1978: Plume model validation. *EA-917-SY*, Workshop WS-78-99, EPRI, Palo Alto, Calif.

Environmental Protection Agency, 1977: *Conference on Air Quality Modeling*. Acme Reporting Company, Washington, D.C.

——, 1978: Guideline on air quality models. *EPA 450/2-78-027, OAQPS No. 1-2-080*, USEPA, Research Triangle Park, N.C., 48 pp.

Fox, D. G., and J. E. Fairobent, 1981: NCAQ panel examines uses and limitations of air quality models. *Bull. Am. Meteorol. Soc.*, **62**, 218–221.

Gifford, F. A., 1972: The form of the frequency distribution of air pollution concentrations. *Proceedings, Symposium on Statistical Aspects of Air Quality Data, Chapel Hill, 9–10 November 1972*, USEPA, Research Triangle Park, N.C., pp. 3-1–3-7.

Hanna, S. R., G. A. Briggs, J. Deardorff, B. A. Egan, F. A. Gifford, and F. Pasquill, 1977: AMS workshop on stability classification schemes and sigma curves—Summary of recommendations. *Bull. Am. Meteorol. Soc.*, **58**, 1305–1309.

Hayes, S. R., 1979: Performance measures and standards for air quality simulation models. *EPA-450/4-79-032*, USEPA, Research Triangle Park, N.C., 160 pp.

Hillyer, M. J., S. D. Reynolds, and P. M. Roth, 1979: Procedures for evaluating the performance of air quality simulation models. *EPA-450/4-79-033*, USEPA, Research Triangle Park, N.C., 146 pp.

Holzworth, G. C., 1980: The EPA program for dispersion model development for sources in complex terrain. *Conference Papers, Second Joint Conference on Applications of Air Pollution Meteorology, New Orleans*, AMS, Boston, pp. 465–468.

Javitz, H. S., and R. E. Ruff, 1979: Evaluation of the real-time air quality model using the RAPS data base, Vol. II: Statistical Procedures. SRI International, Inc., Menlo Park, Calif. (Report for EPA Contract 68-02-2770.)

Larsen, R. I., 1971: A mathematical model for relating air quality measurements to air quality standards. *AP-86*, USEPA, Research Triangle Park, N.C.

Londegren, R. J., 1980: Validation of plume models, statistical methods and criteria. TRC Environmental Consultants, Inc., Wethersfield, Conn., 24 pp.

MacKay, K. P., 1980: Proposed rationale for evaluating the performance of air quality simulation models. EPA Staff Rept. (Draft), Research Triangle Park, N.C., 21 pp.

Mills, M. T., R. Caiazza, D. D. Herbert, and D. A. Lynn, 1980: Evaluation of point source dispersion models. Teknekron Research, Inc., Waltham, Mass. (Report No. R-030-EPA-79 to EPA.)

Nappo, C. J., Jr., 1974: A method for evaluating the accuracy of air pollution prediction models. *Preprints, Symposium on Atmospheric Diffusion and Pollution* (*Santa Barbara*), AMS, Boston, pp. 325–329.

Roberts, J. J. (Ed.), 1977: *Report to the USEPA of the Specialists' Conference on the EPA Modeling Guideline*. USEPA, Research Triangle Park, N.C., pp. 111–116.

Venkatram, A., 1979: A note on the measurement and modeling of pollutant concentrations associated with point sources. *Boundary-Layer Meteorol.*, **17**, 523–536.

Weil, J. C., 1979: Applicability of stability classification schemes and associated parameters to dispersion of tall stack plumes in Maryland. *Atmos. Environ.*, **13**, 819–831. ●

Appendix L

On-Site Meteorological Instrumentation Requirements to Characterize Diffusion from Point Sources—A Workshop, 15-17 January 1980, Raleigh, N.C.

G. F. Hoffnagle,[1] M. E. Smith,[2] T. V. Crawford,[3] and T. J. Lockhart[4]

1. Introduction

A workshop entitled On-Site Meteorological Instrumentation Requirements to Characterize Diffusion from Point Sources was held in Raleigh, N. C., on 15-17 January 1980. The workshop was sponsored by the U.S. Environmental Protection Agency (EPA) and arranged by Environmental Research & Technology, Inc. The purpose of the workshop was to provide EPA with a thorough examination of the meteorological instrumentation and data collection requirements needed to characterize airborne dispersion of air contaminants from point sources, and to recommend, based on an expert consensus, specific measurement techniques and accuracies. Secondary purposes of the workshop were 1) to make recommendations to the National Weather Service (NWS) about collecting and archiving meteorological data that would best support air quality dispersion modeling objectives and 2) to make recommendations on standardization of meteorological data reporting and quality assurance programs.

A representative group of 31 scientists drawn from regulatory agencies, industries, universities, research laboratories, and consulting firms participated in the workshop.[5] The workshop was conducted in three panels for the majority of its work, with plenary sessions for coordination, as follows:

1) Panel A, chaired by Maynard Smith, evaluated data requirements, including model input variables, data representativeness, accuracy requirements, averaging and sampling periods, data archiving, and recommendations to the NWS.
2) Panel B, chaired by Todd Crawford, considered the computations and data required for derived variables used to describe turbulence and diffusion in the planetary boundary layer.
3) Panel C, chaired by Tom Lockhart, reviewed measurement techniques for the requisite meteorological variables.

Summary reports from each of these panels have been integrated as the basis of the workshop recommendations.

Meteorological measurements discussed in the workshop are used in point source diffusion modeling and in analysis of air quality monitoring data (e.g., as discussed in the Modeling Guideline (EPA, 1978a) and the Monitoring Guideline (EPA, 1978b)). The workshop recommendations are intended to help the modeler choose the instrumentation and analysis techniques required to measure accurately the variables needed for his model, rather than to help him choose the "best" model. Most variables previously found useful for dispersion modeling were considered and reasonably attainable accuracies were estimated.

The meteorological variables were divided into three priority classes: 1) essential, 2) desirable, and 3) helpful. These classes reflect a judgment about desirable progress in state of the art modeling and should not limit the modeler in selecting the variables for specific applications.

The workshop results de-emphasize relying on the Pasquill-Gifford-Turner (PGT) stability classification as a primary means of characterizing atmospheric diffusion; rather, they emphasize relying upon on-site measurements of turbulence intensity to characterize the dispersion properties of the atmosphere. This is a major step forward from the assumptions of many models currently used in air pollution control evaluation.

Of course, much model testing and evaluation will be required before the PGT classification scheme can be replaced. Those at the workshop, however, felt that modeling could be advanced by using these variables in the priorities assigned. These workshop recommendations do not constitute EPA policy with regard to meteorological monitoring programs.

2. Essential (Priority 1) measurements

The quantities that have been identified as *essential* to further progress in atmospheric dispersion modeling are:

mean horizontal wind speed ($\bar{u}$) and mean wind direction ($\bar{\theta}$) at 10 meters (m),
mean horizontal wind speed and mean wind direction at source height,
harmonic mean wind speed $(1/\bar{u})^{-1}$ between source height and effective plume height,

---

[1] Environmental Research & Technology, Inc., Concord, Mass. 01742.

[2] Meteorological Evaluation Services, Amityville, N.Y. 11701.

[3] Savannah River Laboratory, E. I. duPont de Nemours, Aiken, S.C. 29801.

[4] Meteorology Research, Inc., Altadena, Calif. 91001.

[5] Attendees: S. P. S. Arya, R. Beebe, F. Binkowski, R. Bornstein, G. Briggs, E. Burt, T. Crawford, J. Dicke, B. Egan, P. Finkelstein, D. Fulbright, B. Greene, G. Hoffnagle, T. Horst, E. Hovind, J. Irwin, J. C. Kaimal, T. Lockhart, P. MacCready, D. Mazzarella, J. McTaggart-Cowan, F. Pasquill, W. Petersen, S. SethuRaman, D. Smith, M. Smith, R. Swanson, D. B. Turner, I. Van der Hoven, A. Venkatram, and A. Weber.

0003-0007/81/020255-07$05.75
©1981 American Meteorological Society

standard deviations of horizontal and vertical wind fluctuations (e.g., σ_θ, σ_ϕ) at 10 m,
temperature difference (ΔT) between 2 and 10 m in unobstructed surface layers,
temperature (T) at 2 m,
heights of the convectively mixed layer (h_c) and the mechanically mixed layer (h_m),
average surface roughness length (z_0),
friction velocity (u_*),
bulk Richardson number (Ri_B),
Monin-Obukhov length scale (L),
turbulent heat flux (H_F),
convective velocity scale (w_*),
water vapor content (e_w) between source height and effective plume height (if the water content of source emissions is important).

Some of these quantities are primary—direct averages or statistical evaluations of direct measurements—while others are derived from combinations of the primary quantities.

a. Wind measurements

For horizontal wind speed, the scalar mean ($\bar{u}_s$) and the resultant vector mean ($\bar{u}_v$) should be reported. The harmonic mean ($\bar{u}_h$) is, however, preferred for calculating initial dilution and plume rise. These variables are defined by

$$u_s = \frac{1}{N}\sum_i u_{s_i}$$

$$u_v = \left[\left(\frac{1}{N}\sum_i u_i \right)^2 + \left(\frac{1}{N}\sum_i v_i \right)^2 \right]^{1/2}$$

$$u_h = \left(\frac{1}{N}\sum_i \frac{1}{u_{s_i}} \right)^{-1}$$

Mean wind direction should be reported as the hourly resultant vector direction ($\bar{\theta} = \arctan \bar{v}/\bar{u}$). Satisfactory measurements of wind speed and wind direction can be made with cup anemometers and wind vanes, with wind vanes and front-mounted propeller anemometers, or with two or three propellers mounted on orthogonal axes (three propellers to include vertical wind velocity).

The standard deviations of wind fluctuations (σ_θ, σ_ϕ or σ_u, σ_v, σ_w) should be calculated directly, not by a range estimator. For sufficient accuracy (~5%), the computation should be based on digital or mathematically equivalent analog techniques.

At monitor sites likely to experience intermittent pollutant concentrations, wind statistics based on a 3 min (or shorter) sampling duration for σ_θ should also be gathered. These σ_θ values can then be combined to match pollutant release intervals in order to model time-integrated concentration fields.

Cup or propeller anemometers must be accurate to within 0.2 m/s ±5% of the wind speed, with a start speed of ≤0.5 m/s, and a distance constant (63% recovery) of ≤5 m. Wind vanes must have a resolution of 1° and an accuracy of 5°. Delay distance (50% recovery) must be ≤5 m with a damping ratio of ≥0.4.

Sixty or more samples will estimate hourly means to within 5–10%. Sample averaging time should be 1–5 s, with a response time of 1 s. Three hundred and sixty or more samples will estimate the hourly standard deviation within 5–10%. The need for 360 samples for calculating hourly standard deviations was questioned during postworkshop discussions. The response from Panel C members stated that 360 samples is a conservative number based on substantial experience. In some instances fewer samples would be needed, but there is no general guide for identifying those situations prior to sampling.

Applied techniques for obtaining winds aloft include tracking superpressure balloons and using Doppler sodar systems.

Ascending free balloons provide data on flow trajectories over a finite time interval. Since the balloon does not follow exactly an air parcel when vertical motion is involved, the results must be interpreted cautiously.

The Doppler sodar technique may be used in either the monostatic or the bistatic mode. Measurements of wind components aloft are accurate to about 0.5 m/s. These measurements routinely provide information to heights of 500 m, and sometimes to heights above 1000 m.

b. Temperature measurements

Ambient air temperature should be measured at 2 m, consistent with the World Meteorological Organization (WMO) standards for ambient temperature measurements. It is recognized that some derived parameters are based on temperatures at 10 m, but this value may be calculated from the 2 m temperature and the 2–10 m temperature difference. The ambient temperature measurement should be accurate to 0.5°C, with a resolution of 0.1°C.

Temperature differences should be measured between 2 and 10 m in an unobstructed surface layer. Measurements should be made with an independent probe system that makes no reference to other measurements of ambient temperature or to the standard deviation of temperature fluctuations (σ_T). Properly shielded and aspirated, a temperature probe system can attain 0.1°C accuracy and 0.02°C resolution with the present state of the art. Standards of probe selection, calibration, and maintenance for research efforts can yield 0.05°C accuracy and 0.01°C resolution. The time constant (63% recovery) of the temperature probe should be 1 min, and the sample averaging time should be greater than or equal to half the time constant.

In siting the probes, care must be taken to maintain the integrity of the ambient environment, particularly the surface, in the immediate vicinity of the tower. This is especially important for temperature probes 2 m above the ground. Site preparation usually requires substantial modification of the surface at a monitoring site. Panel C did not make specific recommendations about site preparation techniques to insure representative measurements.

It is recognized that 0.1°C accuracy is insufficient for distinguishing atmospheric stability in some applications. Greater accuracy is needed to enhance the usefulness of temperature difference measurements.

c. *Mixed layer heights*

Mixed layers may be maintained convectively by surface heating or maintained mechanically by wind-generated turbulent mixing. A mixed layer is classified as either convective or mechanical, as one or the other mechanism is dominant. When both mechanisms are active, the buoyancy-driven mixed layer usually dominates.

The convective layer height (h_c) should be measured routinely by a monostatic sodar with vertical beam. Objective rules have been developed to assist the operators in evaluating h_c, and automatic and pattern recognition methods are under development. Sodar resolution is about 10 m, and the useful range varies from 50 to 1000 m. The sampling time constant is typically 10 s. Convective mixed layer heights derived from sodar and from rawinsonde profiles generally agree to within 50 m.

When the mixed layer height exceeds the range of the sodar system, other methods are available. Aircraft equipped with a fast response temperature sensor readily provide temperature profiles to greater heights. Tethersondes may be used in moderate to light wind conditions where they are not an aviation hazard. A slow-rise radiosonde balloon may be used when tracked by double theodolite observations. These tracked balloons also provide data on winds aloft.

Bistatic sodar can be used to derive the mechanically mixed layer height, h_m, in the stable boundary layer. But the method is of limited value because the characteristic value of h_m lies between 0 and 200 m, while the minimum range of the technique is 50 m. The mechanically mixed layer height is best measured with temperature or velocity turbulence data from an instrumented tower, supported by sodar.

d. *Roughness length*

The roughness length (z_0) at a site may be derived from wind speed profile data under neutral atmospheric conditions or estimated from a description of surrounding surface features. Once estimated, the roughness length may be assumed to be fixed. Consequently, only a limited record of wind profile data is needed.

If z_0 is to be derived from wind profile data over relatively flat, homogeneous terrain, wind measurements should be made at three heights between 2 and 20 m, most desirably with 5% wind speed accuracy, 10 s response time, 30 s sample averaging time, and 1 h sampling duration. If, however, the surface is patchy, with large and varied roughness elements, so that it is difficult to determine the proper instrument heights, z_0 can be estimated indirectly and z_0 values should depend upon wind direction. The estimate should take into account the size and distribution of roughness elements. Several methods for estimating z_0 were referenced by Panel B, but none has been shown to be completely satisfactory. Fortunately, dispersion modeling is not very sensitive to z_0. It is adequate to estimate z_0 to within a factor of 2.

e. *Bulk Richardson number, friction velocity, Monin-Obukhov length, heat flux, and convective velocity*

The derived quantities Ri_B, u_*, L, H_F, w_* can be estimated from temperature and wind measurements.

The bulk Richardson number, Ri_B, is derived from measurements at 2 m and 10 m:

$$\mathrm{Ri}_B = \frac{g \cdot (z_{10})}{T_{10}} \frac{(\Delta T + \Gamma)}{(\bar{u}_{10})^2}$$

where g is the acceleration due to gravity, T_{10} (K) is the ambient temperature at 10 m, and Γ is the dry adiabatic lapse rate. This value of Ri_B is representative of the surface boundary layer only.

The Monin-Obukhov length scale, L, may be evaluated indirectly from the value obtained for Ri_B. Assuming the validity of surface layer similarity theory relationships, L can be obtained by iterative approximation to:

$$z/L = \mathrm{Ri}_B \, kF^2/G$$

where F and G represent integral forms of the empirical flux-profile relations (Benoit, 1977; Irwin and Binkowski, 1980). Both F and G are functions of z/z_0, z/L, and z_0/L. Since the bulk Richardson number is derived from measurements at 2 m and 10 m, z in this application is 10 m. This procedure for obtaining L has been compared with direct measurement results and shown to be within a factor of 2.

The surface friction velocity, u_*, is readily calculated once L is derived. By definition, the function F, used in the iterative technique for obtaining L from profile measurements, is the ratio of the mean wind at the upper measurement height (10 m) to the friction velocity. Therefore

$$u_* = \frac{\bar{u}(z_{10})}{F(z_{10}/z_0,\ z_{10}/L,\ z_0/L)}$$

Similarly, the surface heat flux, H_F, is calculated from L and u_*. Using the definition of the Monin-Obukhov length,

$$H_F = u_*^3 \rho C_p T_{10}/(kgl)$$

Finally, the convective velocity scale, w_*, is derived from the heat flux and the convective layer height, h_c, by:

$$w_* = \left(\frac{g}{T_{10}} \frac{H_F}{C_p} h_c \right)^{1/3}$$

Measurements of incoming solar radiation or net radiation can also be used to estimate w_* (Venkatram, 1978; Briggs, 1975).

f. *Water vapor content measurements*

Water vapor content, e_w, is derived in part from measurements of temperature dew point, T_d. Either of two dew-point sensor techniques may be used: the dew cell technique or the cooled-mirror technique. A measurement accuracy of ±1.5°C over the temperature range −3°C to +20°C is adequate. As a rule, the time constant of the system is not critical—1–30 min. For some applications, such as cooling tower fog formation, a short time constant may be required.

The dew cell technique with a resistance thermometer element has a characteristic time constant of 12–15 min; miniature versions with thermistors are much faster. For measurements at source height or above, it is important to measure the corresponding ambient temperature so that the water vapor content of the air may be calculated.

TABLE 1. Measurements required for Priority 1 variables.

Measurement	Height	Accuracy	Resolution	Sample averaging time	Response Time
u*	10 m, source, plume	0.2 m/s ± 5% 0.5 m/s	— —	1–5 s —	distance const. ≤5 m —
θ*, ϕ*	10 m, source	absolute 5° relative 1°	—	1–5 s	distance const. ≤5 m damping ratio ≥0.4
ΔT	2–10 m	0.1°C	0.02°C	30 s	1 min
h_c	50–2000 m	50 m	10 m	—	10 s
h_m	0–200 m	10–50 m	10 m	—	10 s
T_d**	source, plume,	1.5°C	0.1°C	≥1/2 of response time	1–30 min
T	2 m, source, plume,	0.5°C	0.1°C	30 s	1 min

* Or equivalent wind measurements u, v, w. Requirements specified are for threshold speeds ≤0.5 m/s.

** Required only when cooling tower plumes are modeled.

g. Summary

The basic measurements required for Priority 1 quantities are presented in Table 1 with information on instrument heights, accuracy, resolution, sampling times, and response time.

All the derived quantities depend, ultimately, on the primary measured quantities; some derived quantities depend on others derived first. Table 2 lists all the Priority 1 quantities, the time intervals for their characterization, and the quantities from which they are determined.

3. Desirable (Priority 2) measurements

In addition to the Priority 1 variables, six other variables were recommended as desirable for more accurate modeling in certain situations:

standard deviation of the reciprocal of the wind speed at source height ($\sigma(1/u)$),
mean wind direction at plume height ($\bar{\theta}_v$),
standard deviation of horizontal and vertical wind fluctuations at source height (σ_θ, σ_ϕ),
temperature gradient from source height to effective plume height (ΔT),
total solar radiation (SR).

a. Wind measurements

The Priority 1 wind measurements at source height will provide the data necessary to estimate $\sigma(1/u)$, σ_θ, and σ_ϕ at source height. Remote sensing or tall tower observations of u and θ are required for $\bar{\theta}$ at plume height.

Computational procedures and sampling duration for $\sigma(1/u)$ are similar to those for σ_θ and σ_ϕ already discussed in Section 2. It is recommended that the standard deviation for $\sigma(1/u)$ be computed from 1 min averages of u to retain only the lower frequency variations in wind speed that affect plume rise.

b. Temperature measurements

For sources of modest effective plume height, the temperature gradient from source height to plume top may be obtained from an instrumented tower, with measurement ac-

TABLE 2. Interdependence of Priority 1 variables.

Symbol	Sampling duration	Needed for derivation	Comments
$\bar{u}_v$, $\bar{u}_s$	1 h	u, θ: 10 m, source	—
$\overline{\theta_v}$	1 h	u, θ: source	—
$\overline{1/u_s}$	1 h	u: source	1 min samples of u
σ_θ, σ_ϕ	1 h	θ: 10 m	—
	3 min	θ: 10 m	Optional, depending on source characteristics
ΔT	1 h	ΔT: 2–10 m	—
h_c	1 h	h_c	—
h_m	1 h	h_m	—
z_0	1 h	u: 2 m, 10 m, 20 m	Derived from wind profile measurements or estimated from site description
Ri_B	1 h	ΔT, $\bar{u}$(10 m), $\bar{T}$(2 m)	—
L	1 h	Ri_B, z_0	—
u_*	1 h	z_0, u (10 m), L	—
H_F	1 h	ΔT, L, $\bar{u}_*$, $\bar{T}$(2 m)	—
w_*	1 h	H_F, $\bar{T}$(2 m), h_c, ΔT	—
e_w	3 h	T_d, T: source	—

curacies of 0.1°C. For tall sources, discrete temperature sounding profiles are accurate enough to resolve major temperature gradients through the layer of plume rise. Such temperature profiles, obtained once per hour, will also help resolve h_m and h_s.

c. Solar radiation

Global sun and sky radiation should be measured with pyranometers with 50 W/m^2 accuracy, a resolution of 10 W/m^2, and a time constant (63% recovery) of about 5 s. A complementary net radiometer may have a response time of between 0.5 and 1.0 min. Radiation measurements may be used to estimate w_*.

d. Summary

Six additional modeling variables are desirable. Their derivations require measurements of winds aloft (at plume height), temperature profiles of plume heights, and solar radiation (SR) measurements at the surface. Table 3 lists the Priority 2 variables, the time interval for characterization, and the quantities from which they are immediately determined.

4. Helpful (Priority 3) measurements

Several further variables are helpful when modeling in complex terrain or when considering long-range transport. These include the mean vertical velocity ($\overline{w}$), the Froude number (Fr), the geostrophic wind (U_g), and the state of the ground over which the plume travels. $\overline{w}$ requires measurements of vertical velocity and Fr requires vertical profiles of temperature and velocity to heights exceeding local terrain elevation. Geostrophic winds should be calculated from three-hourly NWS observations near the site. Table 4 summarizes the Priority 3 variables.

5. Recommendations for the Office of the Federal Coordinator for Meteorology

Translating the recommendations of this workshop into specific requirements for new source evaluations would be very costly. It is suggested that the National Weather Service assist as much as possible by altering, where feasible, its current data acquisition and processing procedures and instrumentation.

In the list of recommendations, Panel A has suggested only those items for which major needs exist, and that would have considerable transfer value from NWS sites to pollutant source locations.

Recommendations that would be incompatible with aircraft operations have been avoided because many of the NWS stations are at airports. The list is therefore restricted to low-level measurements (~10 m), or measurements from remote sensors.

Special efforts should be made to obtain more frequent radiosonde data. A series of at least four per day for periods of approximately three to four months in various sections of the United States would be very helpful. They would be of maximum benefit if the program were conducted on a firm schedule that was published well in advance.

The list of recommendations is presented in Tables 5, 6 and 7. The tables detail recommended low-level quantities, low-level measurements desirable at locations relatively free of wakes and runway surfaces, and recommended high-level measurements from remote sensors.

6. Archiving and retrieval of data

Most primary data will be collected by an automated system and will be initially formatted autographically and/or dig-

TABLE 3. Dependence of Priority 2 variables on measurements.

Quantity	Sampling duration	Derivation	Comments
$\sigma(1/u)$	1 h	u: source	1 min averages of u
$\bar{\theta}_s$	1 h	u, θ: plume	—
σ_θ, σ_ϕ	1 h	θ, ϕ: source	—
	3 min	θ, ϕ: source	Optional, depending on source characteristics
ΔT	1 h	ΔT: source-plume	Vertical profile
SR	1 h	Q^*: 1 m	Total solar and net radiation

TABLE 4. Dependence of Priority 3 variables on measurements.

Quantity	Sampling duration	Derivation	Comments
$\overline{w}$	1 h	w: 10 m	Useful for complex terrain modeling
Fr	1 h	ΔT, u: Profile	Useful for complex terrain modeling
U_g	3 h	NWS p, T	Useful for long-range transport modeling
State of the ground	3 h	Observation	—

TABLE 5. Recommended low-level measurements and derived quantities from NWS sites.

	Type of change		
Quantity	Instrumentation	Data processing	Priority
$\bar{u}$ and $\bar{\theta}$ at 10 m		Alter data processing to provide mean values over the hour	1
$\bar{u}$ at 10 m	Improve sensor to provide lower threshold speed		2
$\bar{\theta}$ at 10 m	Improve sensor to provide better damping (to 0.4 but < 0.65)		2
σ_θ		Compute hourly standard deviations	1
σ_v	Same sensor as for $\bar{u}$	Compute hourly standard deviations	As a substitute for σ_θ if it should be easier
Total solar radiation	Additional equipment required	Mean hourly values	1

TABLE 6. Desirable low-level measurements from NWS sites.

	Type of change		
Quantity	Instrumentation	Data processing	Priority
σ_w	New sensors required for vertical wind	Processing to obtain 1 h standard deviation	2
ΔT(2–10 m) & $\bar{u}$ at 10 m	New ΔT sensors required for bulk Richardson number	Requisite processing for 1 h values	2

Note: These data would be useful if they can be obtained at locations relatively free of aircraft wakes, building effects, runway surfaces, etc.

itally, depending on site instrumentation. Autographic data should be reduced to digital form to eliminate difficulties inherent in their archiving, retrieval, and use. In order for data to be available to users, a responsive data archiving and retrieval system is needed.

The following actions are suggested:

1) Fix the retrieval and user requirements for data quantities, including both derived and measured quantities. (There is no point, for example, in archiving data that lose their utility beyond the real-time application.)
2) Establish the data base access strategy to optimize file design.
3) After quantities to be archived are determined, choose a format that accommodates all data, is site-independent, and is compatible with the data base access strategy.
4) Preprocess data on-site as completely as possible, but keep enough detail to permit subsequent processing. Retain complete details for periods of special interest.
5) Establish quality assurance guidelines for preprocessing and on-site data collection processes. Prearchival processing could include flagging of suspect data to warn the user.
6) Maintain site and instrument descriptions, and documentation of processing techniques for each site as an integral part of the system.

TABLE 7. Recommended high-level measurements from remote sensors at NWS sites.

Quantity	Comments	Priority
$\bar{u}$ above 100 m	An elevated wind speed measurement at some standard (agreed-upon) height	1
h_c	The depth of the convective mixing layer, as available from acoustic or other sounders	2
$d\theta/dz$	Measurements of wind shear in the first 2000 m would be helpful	3

Acknowledgments. This workshop was sponsored by the U.S. Environmental Protection Agency under Contract 68-02-3282, administered by D. Bruce Turner, John Irwin, and William Petersen of the EPA Division of Meteorology.

We thank Ms. Margit Kohl for her logistical support of the workshop, and Dr. Arthur Bass and David Strimaitis for their help in preparing and editing the reports.

We give special thanks to Dr. Frank Pasquill, a consultant to ERT, for his active participation in the workshop and his assistance in reviewing and distilling the workshop findings.

References

Benoit, R., 1977: On the integral of the surface layer profile-gradient functions. *J. Appl. Meteorol.,* **16,** 859–860.

Briggs, G. A., 1975: Plume rise predictions. *Lectures on Air Pollution and Environmental Impact Analysis* (*Boston*), AMS, Boston, pp. 59–111.

EPA, 1978a: *Guideline on air quality models. EPA-540/2-78-027,* Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Pk., N.C.

——, 1978b: *Ambient monitoring guidelines for Prevention of Significant Deterioration* (*PSD*). *EPA-450/22-78-019,* Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Pk., N.C.

Irwin, J., and F. Binkowski, 1980: Estimation of the Monin-Obukhov scaling length using on-site instrumentation. In *Atmospheric Environment* (in press).

Venkatram, A., 1978: Estimating the convective velocity scale for diffusion applications. *Boundary Layer Meteorol.*, **15**, 447–452.

●